

**Aus dem Institut für Medizinische Soziologie und Sozialmedizin
des Fachbereichs Medizin der Philipps-Universität Marburg
in Zusammenarbeit mit dem Universitätsklinikum
Gießen und Marburg GmbH Standort Marburg**

**Die Messung des Einflusses von Heterogenität
bei der Überlebenszeitanalyse
von fertilen und subfertilen Männern**

**Eine Bewertung der Heterogenitätsintensitäten
mit Hilfe von parametrischen Frailty-Modellen**

**Inaugural-Dissertation zur Erlangung des Doktorgrades der
Humanbiologie
dem Fachbereich Medizin**

vorgelegt von

**Ronny Westerman
aus Rostock
Marburg/Lahn 2011**

**Angenommen vom Fachbereich Medizin
der Philipps-Universität Marburg am: 17.11.2011**

Gedruckt mit der Genehmigung des Fachbereichs

Dekan: Prof. Dr. Matthias Rothmund

Referent: Prof. Dr. Ulrich Mueller

Koreferent: Prof. Dr. Helmut Schäfer

Danksagungen

An dieser Stelle möchte ich mich bei allen beteiligten Personen bedanken, die mich bei meinem Promotionsvorhaben unterstützt haben.

Besonderer Dank gilt dabei:

Für die kritischen Bemerkungen sowie die Motivation durch Herrn Prof. Dr. Mueller, die mich gerade in meinen Vorhaben gestärkt haben.

Für die Inspiration und den hilfreichen Kommentaren durch
Herrn Andreas Wienke, Institut für Medizinische Epidemiologie, Biometrie und
Informatik der Martin-Luther-Universität Halle-Wittenberg
Herrn Alan Gagnon, Department of Sociology, University of Western Ontario in
London, Canada

Für die Unterstützung bei der Erweiterung des Datensatzes und der verbrachten
Zeit im Einwohnermeldeamt durch
Katharina Belting und Hanna Seydel

Für den regelmäßigen fachlichen Austausch und den heiteren Momenten am
Institut durch
Christian Jacke und Markus Thiede

Sowie den übrigen Mitarbeitern des Instituts für Medizinische Soziologie und
Sozialmedizin, die mir insbesondere in der Endphase den Rücken freigehalten
haben.

Summary

The present thesis will discuss the effect of unobserved heterogeneity in the context of modelling the association between fertility and the post-reproductive life span of males.

Unobserved heterogeneity must eventually be taken into account for successive data analyses and might be useful for the explanation of unexpected results, or for alternative interpretations.

In the worst case one can detect the mortality cross-over for the hazard functions.

The influence of unobserved covariates in a proportional hazard model can be treated by a positive latent random variable, the frailty Z . The frailty concept implies a mixture of individuals in populations varying in their susceptibility to common risks.

In homogeneous populations the frailty variance σ^2 is small, the value for the frailty Z converges to 1.

But when σ^2 increases the frailty variable Z becomes more relevant for affecting the individual hazard intensively by unobserved heterogeneity.

The unexplained heterogeneity can be unshared or shared among individuals. If the individuals or groups of individuals share, then the frailty the individuals risk is common to all group members.

The frailty concept requires, for the parametric paradigm, the specification of one statistical distribution. The most popular parametric specification for the frailty variance follows the gamma distribution. This is one of the most flexible statistical distributions and can be used as an approximation for any other parametric version. It must be mentioned, that there are no biological or empirical arguments justifying the use of the gamma distribution, its simply computational or mathematical convenience that determines the preference of any parametric version for the frailty.

The dataset contains 1408 patients born before 1942 who attended the fertility and sterility office of the department of andrology at Marburg University Hospital for semen analysis between 1949 and 1985.

The assignment into subgroups was carried out by the analysis of semen samples on the basis of medical records of sperm concentrations according to the WHO (2010) classification: to infertile at sperm concentrations of $<15 \cdot 10^6$ per mL, and to fertile men at $\geq 15 \cdot 10^6$ per mL.

The patients can also be classified into three groups: azoospermics, with none sperms in ejaculate; in oligozoospermics, with less than 15 Mio. per mL but more than zero sperms in ejaculate; and in normozoospermics with more than 15 Mio. sperms per mL.

Regarding to the findings of the statistical analysis, we conclude that the fertility status of males is not a powerful predictor for differences in survival.

Using numerical and graphical tests we showed, that the proportional hazard assumption is not violated in any PH-Model.

Hence the impact of the frailty variance could not be verified in the unshared parametric regression models.

In general, frailty models, as well as all other parametric models are afflicted by the problem of finding the right specification for the most informative statistical distribution.

In spite of these drawbacks frailty models perform as a sophisticated tool for the analysis of event history data, especially in epidemiology, demography, biology and other life sciences.

They can be easily used as alternatives to the traditional survival models.

Inhaltsverzeichnis

Inhaltsverzeichnis.....	I
Glossar.....	IX
Tabellen- und Abbildungsverzeichnis.....	XI
1. Einleitung	1
1.1 Hintergrund und Ziele der Studie.....	1
1.2 Die Bedeutung von Frailty-Modellen in den Lebenswissenschaften.....	5
1.2.1 Die Anwendung eines Frailty-Modells in der botanischen Epidemiologie.....	6
1.2.1 Die Anwendung eines Frailty-Modells in der Tierökologie.....	7
1.2.2 Die Anwendung eines Frailty-Modells in der Linkage-Analyse.....	8
1.3 Stand der der Forschung.....	10
1.3.1 Evolutionsbiologische Theorien.....	10
1.3.1.1 Die Akkumulation von Mutationen.....	10
1.3.1.2 Die antagonistische Pleiotrophie.....	11
1.3.1.3 Die Disposable- Soma- Theorie.....	11
1.3.2 Die Relevanz von sozialen Netzwerken bei physischer Gesundheit und Sterblichkeit.....	13
1.3.3 Wechselwirkungen von evolutionsbiologischen Ansätzen und Theorien der sozialen Netzwerke.....	17

1.3.4	Statistischer Zusammenhang von Fertilität und Lebenserwartung bei historischen Populationen.....	23
1.3.4.1	Positive Zusammenhang von Fertilität und Postreproduktiver Lebenserwartung.....	23
1.3.4.2	Negativer Zusammenhang von Fertilität und Postreproduktiver Lebenserwartung	30
1.3.4.3	Statistisch nichtsignifikanter Zusammenhang von Fertilität und postreproduktiver Lebenserwartung.....	33
1.3.5	Zwischenfazit.....	36
1.3.6	Zusammenhang von Fertilität und postreproduktiver Sterblichkeit bei Männern mit Spermiogramminformationen...	46
1.4	Anatomische und physiologische Grundlagen der Spermiogenese.....	52
1.4.1	Anatomie des männlichen Genitaltrakts.....	52
1.4.2	Regulierung der Sexualhormonsynthese.....	53
1.4.3	Funktion des Hodens.....	54
1.4.4	Männliche Infertilität.....	55
1.5.	Beschreibung einiger andrologischen Erkrankungen...	58
1.5.1	Nebenhodenentzündung (Epididymitis).....	58
1.5.2	Hodenentzündung.....	59
1.5.3	Urogenitaltuberkulose.....	59
1.5.4	Varikozele.....	60

1.5.5	Hypospadie.....	61
1.5.6	Kryptochismus.....	62
1.5.7	Hodenkrebs.....	64
1.6.	Analyse der Spermogramm-Parameter.....	66
1.6.1	Einfluss des Alters auf die Spermienkonzentration und Spermienqualität.....	67
1.6.2	Statistische Abgrenzung der Fertilitätsklassifikationen.....	69
2.	Grundbegriffe der Survivalanalyse.....	73
2.1	Definition von Lebensdauer.....	73
2.2	Nichtparametrische Verfahren.....	77
2.3.	Proportional-Hazards-Modell nach Cox.....	80
2.4.	Parametrische Verfahren.....	84
2.4.1	Exponential-Verteilung.....	85
2.4.2	Weibull-Verteilung.....	86
2.4.3	Gompertz-Verteilung.....	87
2.4.4	Gamma-Verteilung.....	88
3.	Einfluss von unbeobachteter Heterogenität in der Mortalitätsforschung.....	90
3.1	Die Wirkung von unbeobachteter Heterogenität an einem rechnerischen Beispiel.....	93
3.2	Die Identifikation eines Mortality- Crossovers.....	97

4. Die Statistische Überprüfung von unbeobachteter

Heterogenität.....	101
4.1 Das Problem der unbeobachteten Heterogenität.....	101
4.2 Numerische Tests für die Proportional-Hazards-Bedingung..	104
4.3 Die graphische Überprüfung der Proportional-Hazards-Bedingung.....	107
4.4 Der theoretischer Hintergrund der Frailty-Modelle.....	110
4.5 Die statistischen Verteilungsfunktionen für den Frailty.....	113
4.5.1 Gamma-Frailty.....	113
4.5.2 Inverse Gauß-Frailty.....	115
4.6 Zielsetzung aller Frailty-Modelle.....	116
4.7 Vollparametrische Frailty-Modelle.....	118

5. Daten- und Projektbeschreibung..... 121

5.1 Forschungsleitende Hypothesen.....	121
5.2 Projektbeschreibung.....	122
5.3 Datengrundlage.....	122
5.4 Patientenakten.....	123
5.5 Vollständigkeit und Konsistenz der Patientenakten...	124
5.6 Recherche und Überprüfung des Vitalstatus.....	126
5.7 Die aktuelle Untersuchungspopulation.....	127
5.8 Definition von Lebenszeit	130
5.9 Aufstellung aller verwendeter Variablen.....	133

5.10	Definition der verwendeten statistischen Parameter und Prüfgrößen.....	135
5.11	Statistik-Software.....	136
6.	Ergebnisse.....	137
6.1	Zusammensetzung des verwendeten Datensatz.....	137
6.2	Durchschnittliche Lebensdauer nach Fertilitätsstatus.....	144
6.3	Zusammenhang von Fertilität und Sterblichkeit.....	146
6.3.1	Ergebnisse der nicht-parametrischen Schätzungen.....	146
6.3.1.1	Vergleich der Survivalverläufe von fertilen und subfertilen Fällen.....	147
6.3.1.2	Vergleich der Survivalverläufe von azoospermen, oligozoospermen und normozoospermen Fällen.....	149
6.3.2	Ergebnisse der Cox-Modelle.....	151
6.3.3	Ergebnisse der alternativen Cox-Modelle.....	158
6.3.4	Parametrische Regressionsmodelle mit Unshared-Frailty...	159
6.4	Einfluss von Lost-to-Follow-up.....	163
6.5	Wohnortspezifische Unterschiede in der Spermienkonzentration.....	165
7.	Diskussion.....	168
7.1.	Datenqualität.....	168
7.1.1	Selektionseffekt nach Wohnort.....	168
7.1.2	Selektionseffekt nach Auffindbarkeit.....	170

7.2	Definition von Lebensdauer.....	174
7.3	Die Auswirkungen der betrachteten Hypothesen auf die Wahl des Studiendesigns.....	176
7.4	Bewertung der Regressionsschätzungen mit dem Faktor Gesamtfertilität.....	180
7.5	Diskussion der Ergebnisse.....	181
7.5.1	Ergebnisse der nichtparametrischen Verfahren.....	182
7.5.2	Semiparametrische Schätzverfahren (Cox-Regression)...	183
7.5.3	Ergebnisse der parametrischen Regressionsmodelle mit Frailty-Varianz	184
7.6	Methodische Probleme der Frailty-Ansätze.....	186
7.6.1	Identifikationsproblem der Proportional-Hazards- Annahme.....	187
7.6.2	Das Spezifikationsproblem parametrischer Frailty-Modelle.....	193
7.6.3	Die Frailty-Varianz als alleiniger Indikator für unbeobachtete Heterogenität.....	194
7.6.4	Die Modellvoraussetzungen für Shared-Frailty-Ansätze...	195
7.7	Alternative Frailty-Modelle.....	198
7.7.1	Correlated-Frailty-Modelle.....	199
7.7.2	Semi-parametrische Frailty-Modelle mit EM-Algorithmus...	199
7.7.3	Penalized Likelihood-Methode (PPL).....	200
7.7.4	Das Frailty-Konzept bei seltenen Ereignissen.....	201

7.8	Mögliche Erklärungsfaktoren für unbeobachtete Heterogenität.....	202
7.8.1	Der Zusammenhang von Rauchverhalten und Unfruchtbarkeit bei Männern.....	203
7.8.2	Äußere Umweltfaktoren und männliche Unfruchtbarkeit...	206
7.8.3	Genetische Komponenten der männlichen Unfruchtbarkeit..	208
8.	Zusammenfassung.....	213
9.	Fazit und Ausblick.....	217
10.	Literaturverzeichnis.....	223
11.	Anhang.....	271

Glossar

Symbole und Abkürzungen

AFT	Accelerated Failure Time Modell
β	Regressionskoeffizient
c	Konstante
Coef.	Regressionskoeffizient
CI	Konfidenzintervall
D	Sterbefälle
d.f.	Freiheitsgrade (degrees of freedom)
E(X)	Erwartungswert
f	Dichtefunktion
F	Kumulative Dichtefunktion
$g(X)$	Vektor für die Kovariablen
HR	Hazard-Ratio
λ	(Baseline) Hazardfunktion
Λ	Kumulative Hazardfunktion
L	Likelihoodfunktion
ϕ	Beschleunigungsfaktor
PLL	Partielle Likelihoodfunktion
P(A)	Wahrscheinlichkeit von Ereignis A
PH	Proportional-Hazard
PHGLM	SAS-Prozedur für die Berechnung von Proportional- Hazard- Modellen
R(t)	Risikomenge zum Zeitpunkt t
S	Survivalfunktion
SE	Standardfehler (Standard Error)

T_b

$(\ln(T) \sim N(m, s^2))$

$T \sim \text{Exp}(\lambda)$

$T \sim \text{Weibull}(a, b)$

$T \sim \text{Gompertz}(a, b)$

$T \sim \Gamma(k, \lambda)$

θ

$\text{Var}(X)$

WHO

Z

Baseline-Zeit

Log-normal-Verteilung mit den
Parameter m und s^2

Exponentialverteilung mit Parameter

Weibullverteilung mit Parameter a, b

Gompertzverteilung mit Parameter a, b

Gamaverteilung mit Parameter k, λ

Theta, Prüfgröße für die Frailty-Varianz
Varianz

World Health Organisation

Zufallsvariable für den Frailty

- Abbildung 1: Bedeutung des männlichen Faktors in Fertilitätsbestimmung
- Abbildung 2: Darstellung der kumulierten Sterblichkeit für Raucher und Nichtraucher
- Abbildung 3: Histogramm Geburtsjahr der Fälle
- Abbildung 4: Histogramm Sterbejahr der Patienten
- Abbildung 5: Histogramm Jahr der ersten Untersuchung
- Abbildung 6: Histogramm Alter zum Zeitpunkt der ersten Untersuchung
- Abbildung 7: Histogramm Spermienkonzentration (Mio./ml) für alle Patienten
- Abbildung 8: Histogramm Spermienkonzentration der oligozoospermen Fälle
- Abbildung 9: Kaplan-Meier- Survival-Funktion für subfertile und fertile Fälle
- Abbildung 10: Kumulierte Hazards nach Nelson-Aalen für subfertile und fertile Fälle
- Abbildung 11: Kaplan-Meier- Survival-Funktion für azoosperme, oligozoosperme, und normozoosperme Fälle
- Abbildung 12: Kumulierte Hazards nach Nelson-Aalen für azoosperme, oligozoosperme und normozoosperme Fälle
- Abbildung 13: Überprüfung der Proportional-Hazards-Bedingung für fertile und subfertile Fälle
- Abbildung 14: Überprüfung der Proportional-Hazards-Bedingung für azoosperme, oligozoosperme und normozoosperme Fälle

- Tabelle 1: Ursachen der männlichen Unfruchtbarkeit
- Tabelle 2: Untere Referenzwerte für die Spermienparameter WHO (2010)
- Tabelle 3: Statistische Abgrenzung von Fertilität und Subfertilität
(Guzick, et al., 2001)
- Tabelle 4: Die Wirkung von Selektion in unterschiedlich
Zusammengesetzten Teilpopulationen anhand eines
fiktiven Beispiels
- Tabelle 5: Aufstellung aller verwendeter Variablen
- Tabelle 6: Vitaldaten der Untersuchungspopulation
- Tabelle 7: Individualdaten und Fertilitätsklassifikation aus den
Angaben der Patientenakten
- Tabelle 8: Restricted Mean für die Lebensdauer insgesamt nach
Fertilitätsstatus
- Tabelle 9: Teststatistik auf Unterschiede im Survivalverlauf
für subfertile und fertile Fälle
- Tabelle 10: Teststatistik auf Unterschiede im Survivalverlauf für
azoosperme, oligozoosperme und normozoosperme Fälle
- Tabelle 11: Cox- Modell zum Einfluss der Spermienkonzentration
auf den Hazard mit Geburtsjahr und Wohnort

- Tabelle 12: Cox- Modell zum Einfluss des Fertilitätsstatus auf den Hazard für subfertile und fertile Fälle
- Tabelle 13: Cox-Modell zum Einfluss des Fertilitätsstatus auf den Hazard für azoosperme, oligozoosperme und normozoosperme Fälle
- Tabelle 14: Scoretest für die Skalierten Schoenfeld-Residuen für subfertile und fertile Fälle
- Tabelle 15: Scoretest für die Skalierten Schoenfeld Residuen für azoosperme, oligozoosperme und normozoosperme Fälle
- Tabelle 16: Alternatives Cox-Modell zum Einfluss des Fertilitätsstatus auf den Hazard für subfertile und fertile Fälle
- Tabelle 17: Alternatives Cox-Modell zum Einfluss des Fertilitätsstatus auf den Hazard für azoosperme, oligozoosperme und normozoosperme Fälle
- Tabelle 18: Exponential-Modell mit Unshared (Gamma) Frailty für subfertile und fertile Fälle
- Tabelle 19: Exponential-Modell mit Unshared (Gamma) Frailty für azoosperme, oligozoosperme und normozoosperme Fälle
- Tabelle 20: Weibull-Modell für subfertile und fertile Fälle
- Tabelle 21: Weibull-Modell für azoosperme, oligozoosperme und normozoosperme Fälle
- Tabelle 22: Gompertz-Modell mit Unshared (Gamma) Frailty für fertile und subfertile Fälle

- Tabelle 23: Gompertz-Modell mit Unshared (Gamma) Frailty
für azoosperme, oligozoosperme und normozoosperme Fälle
- Tabelle 24: AIC für subfertile und fertile Fälle
- Tabelle 25: AIC für azoosperme, oligozoosperme und normozoosperme
Fälle
- Tabelle 26: Mann-Whitney-Test für die Verteilung
der Spermienkonzentration
(normozoosperm und oligozoosperm) nach Wohnort
- Tabelle 27: Mann-Whitney-Test für die Verteilung der
Spermienkonzentration (normozoosperm) nach Wohnort
- Tabelle 28: Reference for Genetic Risks associated with Common
Male Infertility
- Tabelle 29: Mit Infertilität verbundene genetische Störungen

Anhang

- Abbildung A1: Bimodale Population-Hazards und Hazard-Funktion separat
für den jeweiligen Peak, Haugen, et. al. 2009
oligozoosperme und normozoosperme Fälle mit
Untersuchungsjahr und Wohnort

1. Einleitung

1.1 Hintergrund und Ziele der Studie

In einer bereits verfassten Arbeit zur Lebenszeitmortalität von Männern mit unterschiedlichen Fertilitätszuständen mit der identischen Studienpopulation konnte für subfertile Männer ein 1,6-fach höheres Sterberisiko nachgewiesen werden (Groos, 2005). Es stellt sich dabei die Frage, ob dieser messbare Unterschied in der Lebensdauer nur ein Artefakt ist oder ob dieses Phänomen real ist. Bisher konnte ein solcher Zusammenhang für Männer nachgewiesen werden, welche Geburtsjahrgängen angehörten, die einen Zeitraum vor dem Zweiten Weltkrieg einschlossen. Kann ein ähnlicher Sterblichkeitsunterschied zwischen subfertilen und fertilen Männern auch bei einer Erweiterung des Datensatzes ermittelt werden?

Haben dabei infertile Männer immer eine höhere Sterblichkeit oder gelten diese Gesetzmäßigkeiten nur für bestimmte Altersabschnitte?

Die Relevanz von Heterogenität für die Mortalitätsforschung zeigen Vaupel et al. (2006) anhand einer sich seit den 80er Jahren des letzten Jahrhundert abschwächenden Verbesserung der Lebenserwartung in der US-Bevölkerung und einer daraus folgenden zunehmenden Distanz zu Ländern wie beispielsweise Japan, bei denen nach wie vor ein starker Zugewinn an Lebenszeit zu beobachten ist.

Diese Entwicklung ist vermutlich auf die Veränderung der Zusammensetzung der Teilpopulationen in der US-Bevölkerung zurückführen und konnte mit direkt beobachtbaren sozioökonomischen Eigenschaften auch nachgewiesen werden.

Dennoch können Unterschiede in der Sterblichkeit nicht immer durch scheinbar leicht erfassbare Kontrollgrößen wie Alter, Geschlecht oder den physisch feststellbaren Gesundheitszustand erklärt werden. Oftmals existieren andere Faktoren, die einen erheblichen Einfluss auf die Ausfallneigung von Individuen haben, aber nicht direkt beobachtet werden können.

Je intensiver der Effekt von unbeobachteter zufälliger Heterogenität ist, desto wahrscheinlicher sind auch die Parameterschätzungen verzerrt, und die Interpretation dieser Schätzergebnisse ist oftmals sehr schwierig. Das Einwirken der inneren Selektion in der Untersuchungsgruppe der Hochrisikofälle kann entsprechend intensiver sein, sodass sich diese Individuen zu einer Teilgruppe entwickelt haben, in der die Lebensrisiken nun deutlich geringer sind.

Die Lebenserwartung ist für diese Teilgruppe ebenfalls angestiegen. Für den konkreten Fall würde das bedeuten, dass sich die innere Zusammensetzung der Gruppe der Subfertilen im Laufe der Beobachtungszeit verändert hat, während der zunächst jene Fälle herausselektiert wurden, die bereits zu Studienbeginn eine höhere Sterbeneigung hatten, und die nun somit zum größten Teil Mitglieder enthält, die eine deutlich geringere Sterbeneigung aufweisen. Die Selektion in der Gruppe der fertilen Probanden könnte sich vollkommen entgegengesetzt verhalten.

Bedingt durch die scheinbar besseren Lebensrisiken – z.B. ist die Anzahl der Vorerkrankungen deutlich geringer –, könnte die Selektion zu Beginn vergleichsweise weniger intensiv sein, sodass sie sich auf einen späteren Zeitraum verschieben müsste, weil im Laufe der Beobachtungszeit die Zahl der Risikofälle altersbedingt anstiege und somit auch die Mortalität zunähme.

Es gilt somit auch zu klären, ob die unbeobachtete Heterogenität und die damit verbundene Selektion sich auf die gesamte Lebensspanne oder nur auf einzelne Teilzeiträume beziehen, und ob sich die Survivalverläufe zwischen den Vergleichsgruppen unterscheiden.

Ziel dieser Arbeit ist ein Vergleich der Hazard- und Survivalfunktionen von fertilen und subfertilen Männern. Dabei sollen die Einflüsse von Kovariablen wie Alter, Kohortenvarianzen oder Wohnort herangezogen werden, um die Sterblichkeitsunterschiede zu erklären. Statistische Zusammenhänge zwischen dem Hazard bzw. dem Survival und den erklärenden Variablen lassen sich methodisch vielfältig durch Proportional-Hazards-Modelle von Cox oder mit Frailty-Modellen berechnen.

Diese Ansätze basieren auf der Annahme der Proportional Hazards, dennoch können sich die Schätzergebnisse der Modelle unterscheiden, besonders dann, wenn die verwendeten Kovariablen keinen ausreichend hohen Erklärungsgehalt liefern und somit mögliche Effekte von unbeobachteter Heterogenität zulassen. Insbesondere können gerade die Berechnungen mit Cox-Modellen und Frailty-Modellen jeweils zu sehr unterschiedlichen, oftmals auch sehr widersprüchlichen Schätzergebnissen führen.

Die Proportional-Hazards-Modelle von Cox versuchen einen Zusammenhang zwischen den Population-Hazards und einer bestimmten Anzahl von Kovariablen herzustellen, welcher nur dann gültig ist, wenn die Untersuchungspopulation bezüglich der zu untersuchenden Eigenschaft, insbesondere der Ausfallneigung, homogen ist.

Der methodische Vorteil des Proportional-Hazard-Ansatzes von Cox liegt zweifelsohne in der Nichtspezifizierung der Baselinefunktion für die individuelle Ausfallneigung, diese Vorgehensweise ist aber nur dann gerechtfertigt, sobald alle Individuen zu einem gegebenen Zeitpunkt die gleiche Ausfallswahrscheinlichkeit besitzen.

Viele epidemiologische Studien zeigen aber auf, dass sich die betrachteten Studienpopulationen aus einer Vielzahl von Individuen zusammensetzen, die sich bezüglich ihrer biologischen Eigenschaften doch stark unterscheiden können.

Insbesondere können sich Anfälligkeiten (Frailties) gegenüber Krankheiten oder dem Tod im Laufe eines Lebens verändern und sind somit nicht konstant. Die Vielseitigkeit in der Anfälligkeits-Neigung können durch Frailty-Modelle beschrieben werden, die eine Parametrisierung der Variabilität des individuellen Ausfallrisikos vornehmen. Die Modellierung dieser Variabilität kann durch halbwegs gutartige stetige Verteilungsfunktionen vorgenommen werden. Dabei ist aber anzumerken, dass die Art der Parametrisierung häufig mit biologischen Erklärungsfaktoren nicht ausreichend erklärt werden kann.

Ziel dieser Arbeit ist der Vergleich und eine Gegenüberstellung der Schätzergebnisse der einzelnen parametrischen und semiparametrischen Schätzverfahren. Ebenso werden die Modellannahmen kritisch diskutiert und es wird ermittelt, ob eine entsprechende Parametrisierung der Frailty-Neigung notwendig ist oder nicht. Die in Groos (2005) berücksichtigte Studienpopulation ist ebenfalls Gegenstand der aktuellen Untersuchung, bei der neben einer Fortsetzung des Mortalitäts-Follow-ups auch noch eine Erweiterung des Datensatzes erfolgt.

1.2 Die Bedeutung von Frailty-Modellen in den Bio- und Lebenswissenschaften

Frailty-Modelle sind sehr vielseitig einsetzbar und lassen sich daher auch bei verschiedenen Fragestellungen in den lebens- und naturwissenschaftlichen Disziplinen wie der Physik, Biologie und Medizin immer dann besonders gut anwenden, wenn unbekannte Einflussfaktoren, die sich nicht direkt oder nur schwer beobachten lassen, potentiell die betrachteten Studienergebnisse verzerren könnten. Zur Veranschaulichung sollen dazu drei Beispiele aus verschiedenen Anwendungsbereichen in Kürze dargestellt werden.

1.2.1 Die Anwendung eines Frailty-Modells in der botanischen Epidemiologie

In der botanischen Epidemiologie stellt sich oftmals das Problem, welche relevanten Faktoren dazu beitragen, dass sich bestimmte Pflanzenkrankheiten (z.B. airborne diseases) in einigen Regionen besonders gut ausbreiten können und warum bestimmte Pflanzenarten häufiger infiziert werden.

Bei der Untersuchung der Ausbreitung von Braun-Rost (*Pucciniales triticina*), eines Schadpilzes bei Weizen, haben zwei Studien (Aylor, 1990; Robert, 2003) gezeigt, dass neben dem Bepflanzungsabstand der einzelnen Weizenpflanzen innerhalb eines Feldes auch die jeweilige individuelle Beschaffenheit der Pflanze verantwortlich für die Verbreitung dieser Krankheit ist. Dabei wird die Intensität des Sporenbefalls einzelner Blätter auch als mögliche Ursache für die Ausbreitung von Braun-Rost angesehen (Soubeyrand et al., 2007).

Aus praktischen Gründen können nicht bei allen Pflanzen die einzelnen Blätter auf ihren individuellen Befall untersucht werden, daher erfolgt die Modellierung der unbeobachteten Varianz der Blattbeschaffenheit mit Hilfe einer Zufallsvariable, der Frailty. Dieser Frailty-Effekt in der Variabilität der Weizenpflanzen konnte bestätigt werden (Soubeyrand et al., 2007); dabei wird vermutet dass weniger robuste Blätter unter Berücksichtigung des Standortes (Licht, Wasserzufuhr) häufiger von Braun-Rost befallen werden.

1.2.2 Die Anwendung eines Frailty-Modells in der Tierökologie

Der Einfluss von unbeobachteter Heterogenität innerhalb spezifischer Populationen wird in vielen Studien oftmals nur unzureichend behandelt, da es oftmals nur unter bestimmten Voraussetzungen möglich ist, die individual-spezifischen Faktoren zu ermitteln. Als eine der wenigen tierökologischen Studien, die auch Effekte von Zufallseinflüssen berücksichtigt haben, ist die Arbeit von Wintrebert et al. (2005) zu nennen, die einen Zusammenhang zwischen dem Survival und den Brutaktivitäten bei Kittiwake (*Rissa tridactyla*) untersucht hat.

Bei der Kittiwake handelt es sich um eine Möwenart, die überwiegend im Nordatlantik und der Arktis verbreitet ist und sich durch eine relativ lange Lebensdauer auszeichnet. Untersucht wurde dabei, inwiefern die Häufigkeit der Brutaktivitäten von der Lebensdauer der Tiere bestimmt wird und welche weiteren Umweltbedingungen den Bruterfolg noch beeinflussen könnten.

Bei der Rekrutierung der Vögel stellte sich heraus, dass sich jüngere Tiere, die noch nie gebrütet haben bzw. bisher nur wenige Brutversuche unternommen haben, mittels der Capture-Recapture-Methode deutlich schwerer fangen ließen als ältere Tiere. Mögliche Unterschiede in der Fangwahrscheinlichkeit könnten daher auch die Studienergebnisse beeinflussen. Die individuelle Varianz in der Bruthäufigkeit wurde durch eine lognormalverteilte Zufallsvariable (Frailty) parametrisch geschätzt.

Ebenso konnte ein positiver Zusammenhang zwischen dem Survival und der Bruthäufigkeit ermittelt werden, wobei zwischen männlichen und weiblichen Tieren keine systematischen Unterschiede festgestellt werden konnten. Die beste Modellanpassung konnte für das Lognormal-logistic-Modell gezeigt werden.

1.2.3 Die Anwendung eines Frailty-Modells in der Linkage-Analyse

Frailty-Modelle sind auch immer dann anzuwenden, wenn genetische Faktoren bei der Vererbung von (seltenen) Krankheiten untersucht werden sollen. Als genetische Komponente für viele Erkrankungen wie z.B. für Brustkrebs oder Alzheimer wird das Alter bei Erkrankungseintritt innerhalb der Familie betrachtet. Die exakten Altersangaben sind häufig aufgrund von Zensierungen und Trunkierungen nicht verfügbar, sodass viele Standardprozeduren zur Linkage-Analyse nur eingeschränkt durchgeführt werden können.

Einen wichtigen wissenschaftlichen Beitrag für dieses Identifizierungsproblem liefert die Linkage-Studie von Jonkers et al. (2009) zu intervallzensierten Migränedaten. Ziel dieser Arbeit war mit Hilfe eines additiven (Gamma-)Frailty-Modells herauszufinden, ob das Alter bei der ersten Migräneattacke in irgendeiner Form vererbbar sei und ob sich dabei entsprechende Gene identifizieren ließen, die mit dem Krankheitsbild verknüpft werden könnten. Die Datengrundlage in dieser Studie umfasste mehr als 4000 Zwillingspaare, wobei nicht bei allen Untersuchungspersonen das Alter bei der ersten Erkrankung bekannt war.

Daher wurde eine entsprechende Frailty-Varianz (Gamma-Spezifikation) definiert, welche die genetischen Effektgrößen für die Geschwisterpaare umschloss.

Der Nachweis einer Vererbung des Alters konnte mit einem relativ hohen Wert von 0,42 erbracht werden, als Lokus für einen Linkage konnte dabei das Chromosom 19 (Lod-Score 1,8) identifiziert werden (Jonkers et al., 2009).

Nach dieser kurzen Einführung in die Thematik und die Anwendungsfelder von Frailty-Modellen sollen zunächst die wichtigsten theoretischen Überlegungen zur Untersuchung eines statistischen Zusammenhangs zwischen der Fertilität und dem Survival herausgearbeitet werden.

1.3. Stand der Forschung

1.3.1 Evolutionsbiologische Theorien

Aus evolutionsbiologischer Sicht hat es viele theoretische Bemühungen gegeben, eine Erklärung für das Zusammenwirken von Fertilität und Lebenserwartung abzubilden. Dabei können im Wesentlichen drei wissenschaftliche Ansätze aufgeführt werden, die durchaus auch in der gegenwärtigen Diskussion noch eine Bedeutung haben.

1.3.1.1 Die Akkumulation von Mutationen

Die Theorie der Akkumulation von Mutationen (Medawar, 1952) im oberen Alter besagt, dass Personen, bei denen sich bereits in frühen Lebensstadien genetische Mutationen klinisch manifestieren, nur geringe oder sogar zu keiner Zeit die Möglichkeit haben sich zu reproduzieren. Treten dagegen diese Mutationen erst im oberen Lebensalter auf, so können sich diese Individuen durchaus reproduzieren, bevor mit gesundheitlichen Einschränkungen durch eine altersbedingte Krankheit zu rechnen ist (Le Bourg, 2001).

Die Theorie von Medawar ist daher kritisch zu bewerten, weil sie davon ausgeht, dass die oberen Altersstufen häufiger von Krankheiten betroffen seien als die jüngeren Altersgruppen, was dann durch die reine Tautologie eines mit zunehmendem Alter ansteigenden Sterberisikos bekräftigt wird.

1.3.1.2 Die antagonistische Pleiotrophie

Gemäß der antagonistischen Pleiotrophie-Theorie (Williams, 1957) sollen bestimmte Gene, welche evolutionär ausselektiert werden, sobald sich die individuelle Fitness im reifen Lebensalter reduziert, für die hohe Fitness im jungen Erwachsenenalter verantwortlich sein. Diese Gene besitzen mehrere unterschiedliche Funktionen, die entsprechend für das jeweilige Lebensalter charakteristisch sind. Sie können sich während der reproduktiven Phase zunächst positiv auf die Gesundheit auswirken, im Laufe des Lebens werden sie dann folglich ausselektiert und haben nun eher negative Eigenschaften, welche dann im oberen Lebensalter kumulieren.

1.3.1.3 Die Disposable-Soma-Theorie

Kirkwood (1993,1999) führt an, dass die bisher dargestellten Zusammenhänge zwar richtig seien, aber wegen des Aspekts des disposable Somas eine notwendige Verminderung der Regenerierungsaktivitäten in den Zellfunktionen bei voraussichtlich zu erwartenden geringen Lebenszeiten mit berücksichtigt werden müsse.

Das bedeutet, dass die Reproduktion sich bei jenen Individuen beschleunigen muss, bei denen sich die Lebensspanne zeitlich verkürzt. Dabei wird von einem allgemeinen Kalkül der Individuen ausgegangen, welche dann entsprechend zwischen Reproduktion und Reaktivierung der Zellfunktionen abwägen müssen. Sobald die Lebensbedingungen für Individuen besonders günstig erscheinen, ist es für diese eher ratsam, mehr Energie und Zeit in die Mechanismen der eigenen Regeneration als in die Aktivitäten der Arterhaltung zu investieren, da sich die Phase der Reproduktion zeitlich verlängert hat und somit sich der Alterungsprozess verlangsamt. Wenn sich die allgemeinen Bedingungen nun deutlich verschlechtern, muss die Reproduktion in einer viel kürzeren Zeitspanne ablaufen, da die individuelle Fitness durch die nun zunehmende Alterung abnimmt. Dieses setzt aber voraus, dass sich alle Individuen zum einen stets rational verhalten und sich entsprechend ihrer Lebenssituation für die richtige Optionen entscheiden und zum anderen auch über die identischen Chancen bzw. über entsprechende Ressourcen für eine ausgewogene Allokation verfügen.

Die drei dargestellten theoretischen Ansätze beschreiben zwar unterschiedliche Mechanismen, dennoch unterstellen sie übereinstimmend, dass eine Verlagerung insbesondere der weiblichen Fortpflanzungsaktivitäten auf ein höheres Lebensalter zu einer Verlängerung der Lebenserwartung bei diesen Individuen führen müsse (Gagnon et al., 2005).

Allerdings lässt sich auch empirisch nachweisen, dass die Häufigkeit von genetischen Mutationen bzw. Körperlichen Missbildungen bei den Nachkommen zunimmt, sobald die Reproduktion auf ein höheres Lebensalter aufgeschoben wird.

Für männliche Organismen scheint ein ähnlicher Zusammenhang zwischen Lebenserwartung und Fertilität aus evolutionsbiologischer Sicht weniger zutreffend zu sein, da die Investitionen in die Qualität der Nachkommen weniger ausgeprägt zu sein scheinen.

1.3.2 Die Relevanz von sozialen Netzwerken bei physischer Gesundheit und Sterblichkeit

Die bisherige Abhandlung der Theoriebildung konzentriert sich zunächst auf rein biologische Ansätze, welche kaum einen Zusammenhang mit der Sozialebene herstellen bzw. auf die Interaktion von Individuen eingehen. Dabei ist empirisch erwiesen, dass gerade der Zugang zu sozialen Netzwerken und den damit verbundenen instrumentellen, informellen und emotionalen Ressourcen (House und Kahn, 1985) sich positiv bzw. negativ auf den Gesundheitszustand, die Reproduktion und daher auch auf die Entwicklung der Lebenserwartung auswirkt. Instrumentelle Ressourcen schließen dabei die materiellen Mittel wie die finanzielle Unterstützung ein, aber auch die physische Hilfe von Fremden bzw. der eigenen Familienangehörigen, die bei der Bewältigung der alltäglichen Lebensaufgaben entsprechend genutzt werden.

Informelle Hilfe grenzt alle Maßnahmen ab, mit denen sich ein Individuum Zugang zu Informationen aneignet, um bestehende, aber auch neu auftretende Probleme eigenständig bewältigen zu können. Emotionale Unterstützung umfasst dabei alle individuellen Einstellungen wie Zuneigung, Empathie, Verantwortungsbewusstsein oder Vertrauen, die ein entsprechendes soziales Gefüge kennzeichnen und ein gesellschaftliches Zusammenleben erst ermöglichen. Gerade Letzteres ist kennzeichnend dafür, warum sich Individuen für oder gegen eine Paarbeziehung mit oder ohne Nachwuchs entscheiden.

Durch relativ enge soziale Verbindungen mit anderen Personen erhalten die Individuen entsprechend Zugang zu für sie lebensnotwendigen Ressourcen, ohne dabei in einem höheren Maß psychischem Stress ausgesetzt zu sein. Somit können soziale Netzwerke in diesem Zusammenhang auch als Stresspuffer angesehen werden (Cohen, 2004). Soziale Netzwerke können nur dann diese Funktion übernehmen, wenn die für ein Individuum benötigten Ressourcen von anderen Personen aus dem Sozialgefüge in ausreichendem Umfang angeboten werden und für alle auch gemeinsam zugänglich sind. Gerade in gesellschaftlichen Krisensituationen wie beispielsweise während der Phase einer wirtschaftlichen Rezession sind gemeinsam zugängliche Ressourcen eher knapp und würden dann eher das Stresspotential in der Bevölkerung erhöhen.

Ebenso ist soziale Unterstützung am effektivsten, wenn sie situationsgebunden stattfindet und eher rational erfolgt. Beispielsweise würde das Verleihen von Geld an eine andere Person in Notsituationen nur dann effektiv sein, wenn diese Person gerade ihre Arbeitsstelle verloren hätte.

Dieselbe Handlung wäre in der Situation völlig irrational, wenn die hilfesuchende Person beispielsweise schon über eigene ausreichende Ressourcen verfügte, diese aber in dieser Situation nicht beanspruchen möchte.

Empirisch konnte die Wirkung sozialer Unterstützung als Stresssensor in einer prospektiven Studie mit schwedischen Männern ab einem Alter von 50 Jahren bei einem vergleichsweise relativ guten körperlichen Allgemeinzustand nachgewiesen werden (Rosengren et al., 1993). Dabei zeigte sich, dass bestimmte Personen, welche bis zum Zeitpunkt des Studienbeginns intensiveren Stresssituationen ausgesetzt waren, anschließend in der siebenjährigen Follow-up-Beobachtung deutlich höhere Sterberisiken aufwiesen als die Vergleichsgruppe von Männern, die nur einer geringfügigen oder überhaupt keiner Stresssituation ausgesetzt waren. Es konnte auch gezeigt werden, dass emotionale Unterstützungen durch andere Personen bei den beobachteten Männern weder eine erhöhte Wahrnehmung noch eine präventive Wirkung auf den psychischen Stress hatten. Genauso erweist sich soziale Integration nicht als geeignete Maßnahme, um Stress bei den Probanden zu reduzieren.

Ein völlig anderes Studienergebnis konnte in einer weiteren prospektiven Studie von Berkman und Syme (1979) bei Bewohnern des Alameda County in Kalifornien nachgewiesen werden.

Für die Ermittlung eines empirischen Zusammenhangs zwischen Mortalität und sozialer Isolation entwickelten die Autoren einen entsprechenden Social Network Index, welcher nicht nur die Häufigkeit der sozialen Kontakte berücksichtigte, sondern auch deren relative Wichtigkeit.

Demzufolge wurden intimere Kontakte wie zu Familienangehörigen und Freunden aus dem direkten Umfeld stärker gewichtet als soziales Engagement in Vereinen oder der Kirche. Abermals konnten die Ergebnisse zeigen, dass bei nur wenigen oder kaum vorhandenen sozialen Kontakten (durch geringe Index-Scores gekennzeichnet), sowohl bei Männern als auch bei Frauen, das Sterberisiko 2,3- bzw. 2,8-mal (Shye et al., 1995) höher war als bei Personen, die deutlich höhere Scorepunkte des Social Network Index erzielten.

In einer weiteren Längsschnittstudie von Giles und Kollegen (Giles et al., 2005) zu den Effekten von sozialen Netzwerken auf die Lebenserwartung älterer australischer Männer und Frauen hatten soziale Kontakte zu Freunden und Bekannten stärkere protektive Wirkungen auf die Lebenserwartung als Beziehungen zu den eigenen Kindern und Verwandten.

Insgesamt ist noch relativ unklar, inwiefern sich die materielle und emotionale Unterstützung durch die bereits erwachsenen Kinder auf die Lebensdauer der Elterngeneration auswirkt.

In Abhängigkeit vom individuellen Gesundheitszustand der Eltern und einem möglichen Transfer von materiellen und nichtmateriellen Ressourcen zwischen den Generationen, bei dem die Kinder eher häufiger Ressourcen nachfragen, als sie entsprechend anbieten können, steigen die direkten und indirekten Kosten für die Eltern, sodass sich dadurch ihre Lebenserwartung reduzieren könnte.

1.3.3 Wechselwirkungen zwischen evolutionsbiologischen Ansätzen und den Theorien der sozialen Netzwerke

Smith et al. (2002) vermuteten spezifische Wechselwirkungen zwischen den evolutionsbiologischen Ansätzen und den Theorien der sozialen Netzwerke und entwickelten daraufhin 7 Hypothesen, welche eine Beziehung zwischen Reproduktion und Lebenserwartung herstellen. Evolutionsbiologische Elemente finden sich in den Hypothesen H1-H3 wieder, dagegen können den Erklärungen von sozialen Mechanismen die Hypothesen H4-H6 zugeordnet werden.

Evolutionenbiologische Konzepte (nach Smith et al., 2002):

- H1: As age at first birth increases, post-reproductive parental longevity increases.
- H2: As the number of children decreases, post-reproductive parental longevity increases.
- H3: As the age at last birth increases, post-reproductive parental longevity increases.

Soziale Mechanismen (nach Smith et al., 2002):

- H4: As age at first birth increases, post-reproductive parental longevity increases.
- H5A: As the number of children increases, post-reproductive longevity increases.
(due to access to greater informal social support)

H5B: As the number of children increases, post-reproductive longevity decreases
(due to wealth flows largely moving from parent to child, children are also high-parity themselves, thereby reducing their supportive capabilities).

H6: As the age at birth increases, post-reproductive parental longevity decreases.

Die evolutionsbiologische Konzepte (Hypothesen H1-H3) unterstellen jeweils in Abhängigkeit von einem hohen Alter bei der ersten Geburt, einer geringeren Kinderzahl und einem niedrigen Alter beim letzten Geburtseignis einen positiven Zusammenhang mit dem postreproduktiven Survival. Die sozialen Mechanismen (Hypothesen H4 – H6) dagegen unterstellen jeweils eine entgegengesetzte Wirkung zwischen dem postreproduktiven Survival und den Erklärungsfaktoren, dem Alter bei der ersten Geburt und der maximal erreichten Kinderzahl.

Deutlich ist dabei zu erkennen, dass einige Hypothesen (H1 und H4 sowie H2 und H5B) zwar die gleichen theoretischen Zusammenhänge unterstellen, aber unterschiedliche Wirkungsmechanismen hervorrufen; andere Ansätze wie H2 und H5B, welche die Parität betreffen, sowie H3 und H6, welche eine Aufschiebung des letzten Geburtseignisses auf ein höheres Lebensalter implizieren, sind eher konträre Handlungsfolgen.

Mögliche Tautologien der dargestellten Hypothesen wurden mit Hilfe der Utah Population Data Base (UPDB) überprüft (Smith et al., 2002).

Die Datengrundlage des UPDB umschließt Individualinformationen von mehr als 7,1 Millionen Personen zu spezifischen Gesundheitsmerkmalen, Inzidenzen von verschiedenen Krebsarten und anderen Krankheiten und bietet daher auch umfassende Möglichkeiten zur Analyse der Vererbbarkeit von genetischen Mutationen.

Für die Analysen wurden 13.897 verheiratete Paare herangezogen, die zwischen 1860 und 1899 eine Eheschließung eingegangen sind.

Da im Bundesstaat Utah eine der größten Gemeinden mormonischen Glaubens existiert, welche auch polygamistische Lebensformen akzeptiert, wurden aufgrund von potentiellen Verzerrungen bei der Analyse jene Ehen ausgeschlossen, die mehr als einen Ehepartner umfassten. Gemäß dieser Studie lassen sich entsprechend folgende Ergebnisse in Kurzform zusammenfassen: Eine Aufschiebung des Alters für das letzte Geburtseignis und eine geringe Parität wirken sich positiv auf die Lebenserwartung der Mütter aus, somit erweisen sich die in H2 und H3 beschriebenen Zusammenhänge als konsistent.

Ein Zusammenhang zwischen einer niedrigeren Kinderzahl und einer höheren Lebenserwartung in der postreproduktiven Phase der Mütter bestätigt damit auch die Richtigkeit der Theorien der antagonistischen Pleiotrophie und des disposablen Soma.

Die Lebenserwartung der Mütter bei hoher Parität reduziert sich deshalb, da bei einer höheren Anzahl von Kindern der Bedarf an verfügbaren Familienressourcen größer ist.

Bei den verfügbaren Ressourcen, die Kinder von ihren Eltern erhalten, handelt es sich um knappe Güter, die nur bei einer geringen Parität effizient eingesetzt werden können.

Zwar können erwachsene Kinder einen Beitrag zur Familienproduktion gemeinsamer Ressourcen leisten und damit den Bereitstellungsdruck knapper Ressourcen auf Seite der Eltern verringern, dieses setzt aber voraus, dass sich die erwachsenen Nachkommen in räumlicher Nähe zum Elternhaus aufhalten.

Erstaunlicherweise besteht kein Zusammenhang zwischen der Aufschiebung des Alters beim letzten Geburtseignis und der Lebenserwartung, deshalb können die Hypothesen H1 und H4 verworfen werden. Eine Verbesserung der Lebenserwartung, bezogen auf geringere Parität und eine aufgeschobene Reproduktion, ist bei verheirateten Männern wesentlich schwächer ausgeprägt als bei Frauen. Insbesondere bei Männern, bei denen die Reproduktion in der Lebensphase erst relativ spät stattfindet, scheint sich dieses späte Geburtseignis verlängernd auf die Lebenserwartung auszuwirken (Mueller, 2004). Begründet wird dies oftmals mit einem langsameren altersspezifischen Sinken der Fekundität und einer sexuell höheren Aktivität im Vergleich zu anderen Männern im gleichen Lebensalter (Smith et al., 2002).

Eine andere Erklärung verweist darauf, dass ältere Väter häufiger mit jüngeren Frauen verheiratet sein müssten. Es ist auch bekannt, dass eine zwischen Ehepaaren bestehende Altersspanne sich positiv auf die Lebenserwartung (Drefahl, 2008, 2010) des Ehemanns auswirkt. Wenn dieser Alterseffekt eliminiert wird, lässt sich zeigen, dass eine aufgeschobene Vaterschaft bei älteren Männern in keinem Zusammenhang mit einer Verlängerung der Lebenserwartung steht.

In einer weiteren wissenschaftlichen Arbeit von Gagnon et al. (2005) wurde eine ähnliche Problemstellung untersucht, die Untersuchungsergebnisse unterschieden sich aber in einigen Punkten von jenen, die Smith et Al. (2002) ermitteln konnten. Es wurden historische Individualdaten des Registre de population du Québec und des Programm de recherche en démographie historique (PRDH) der Universität Montreal genutzt, welche sämtliche verfügbaren Informationen zu Geburtsort, Heirats- und Sterbeereignissen, Namen der Ehepartner und weiteren sozioökonomische Informationen von Bewohnern des Saint-Lawrence Valley im siebzehnten und achtzehnten Jahrhundert beinhalten und eine Stichprobe von 1.950 verheirateten Paaren umfassen.

Im Gegensatz zu den Schätzergebnissen mit der Utah Population konnte ein eindeutiger Beweis eines Zusammenhangs zwischen einer Aufschiebung des ersten Geburtseignisses auf ein höheres Lebensalter und einer Verlängerung der Lebenserwartung nachgewiesen werden, was wiederum für die Richtigkeit der Hypothesen H1 und H4 spricht.

Dieses Ergebnis stützt sich dabei aber auf eine relativ kleine Datengrundlage und würde dann bei einer Erweiterung der der Fallzahl möglicherweise negiert werden.

Beide Studien zeigen auch systematische Unterschiede bei der Betrachtung der Effekte in der männlichen Bevölkerung, bezogen auf den Zusammenhang zwischen Fertilität und Lebenserwartung.

In der Utah-Bevölkerung konnte ein positiver Einfluss von später Fertilität und geringer Parität auf die Lebenserwartung der Ehemänner ermittelt werden, der im Vergleich zur weiblichen Untersuchungsbevölkerung deutlich niedriger ausfiel. Der Nachweis einer ähnlichen Beziehung zwischen den Ereignissen maximal erreichter Geburtenanzahl und der Lebensdauer erweist sich mit Quebec-Daten als inkonsistent.

Gagnon et al. (2005) führen dagegen an, dass bei Männern nicht die maximal erreichbare Zahl von Lebendgeburten entscheidend sei, sondern eher die Anzahl der Kinder, die das 18. Lebensjahr überlebt hätten. Dies bedeute auch, dass biologische Einflüsse der Fertilität auf die Lebenserwartung bei Männern eine geringere Bedeutung haben könnten als bei Frauen.

1.3.4 Statistischer Zusammenhang von Fertilität und Lebenserwartung bei historischen Populationen

Es folgt eine kurze zusammenfassende Darstellung von Studienergebnissen einiger einschlägiger wissenschaftlicher Arbeiten, welche zur Zielsetzung hatten, eine lebenslaufspezifische Assoziation zwischen Fertilität und der postreproduktiven Lebenserwartung zu ermitteln. Dabei ist anzumerken, dass sich diese Studien sowohl in der Methodik der Analyse, der Datenstruktur selbst, den verwendeten Indikatoren zur Quantifizierung von Sterblichkeit und Fertilität sowie daraus folglich auch in den Schätzergebnissen unterscheiden. Dabei zeigt sich eine Variation der Ergebnisse von positiven, negativen sowie nicht signifikanten Zusammenhängen der benannten Untersuchungsgrößen.

1.3.4.1 Positiver Zusammenhang zwischen Fertilität und postreproduktiver Lebenserwartung

Beginnend mit den Studienergebnissen für einen postulierten positiven Zusammenhang von Fertilität und Lebenserwartung kann zunächst die Arbeit von Korpelainen (2000) angeführt werden. Die verwendete Datengrundlage setzt sich aus verschiedenen Teilpopulationen Europas zusammen: dem Großadel und der Aristokratie aus Großbritannien, Frankreich, Spanien, Deutschland, den Niederlanden, Österreich, Russland, Dänemark, Schweden sowie fünf Familien aus den ländlichen Regionen in Ost-Finnland.

Sie enthält außerdem umfangreiche statistische Informationen zum Sterbealter der Eltern, der Anzahl aller Nachkommen und den Überlebenden bis zum Erreichen des Erwachsenenalters, dem Alter der Mutter bei der ersten Geburt, dem Alter der Mutter bei der letzten Geburt und der Zeitspanne der gesamten reproduktiven Phase. Die jeweiligen Elternteile der untersuchten Familien wurden in der Zeitspanne von 1700 bis 1899 geboren und waren alle verheiratet, sodass ein möglicher Effekt durch nichteheliche Geburten ausgeschlossen ist.

Als Indikatoren für gesundheitliche Beeinträchtigung und die Mortalität dienten die individuelle Fitness der Väter und Mütter sowie der Reproduktionserfolg (lifetime reproductive success, LRS), gemessen an der Anzahl der überlebenden Kinder.

Die Studienergebnisse von Korpelainen (2000) konnten erneut morbiditäts- und mortalitätsspezifische Kohorteneffekte aufzeigen, bei denen sich die Sterbealter der Elterngeneration und die Überlebenswahrscheinlichkeit der Nachkommen im Kohortenvergleich unterscheiden und gerade bei jüngeren Geburtsjahrgängen eine deutlich günstigere Entwicklung zu erkennen ist als bei den Angehörigen älterer Kohorten.

Im direkten Subgruppenvergleich hatten Frauen aus bürgerlichen Familien eine durchschnittlich etwas höhere Lebenserwartung als Frauen aus der finnischen Landbevölkerung.

Bei den Männern gab es dagegen keine signifikanten Sterblichkeitsunterschiede zwischen der ländlichen und bürgerlichen Bevölkerungsgruppe. Insgesamt hat sich aber die durchschnittliche Lebenserwartung in beiden Untersuchungsgruppen innerhalb von 100 Jahren deutlich verbessert, im gleichen Zeitraum ist auch die durchschnittliche Kinderzahl gesunken, wobei aber die Überlebenswahrscheinlichkeit der Nachkommen in der bürgerlichen Teilpopulation immer noch höher war.

Etwas überraschend weist die ländliche Bevölkerung, bezogen auf ihre individuelle Fitness, einen etwas besseren allgemeinen Gesundheitszustand auf, welcher vermutlich mit dem späteren Einsetzen der Reproduktion in einem höheren Lebensalter und damit auch einer Verlängerung der Reproduktionsphase verbunden ist.

Wie in der Hypothese H2 von Smith et al. (2002) formuliert wird, wirkt sich ein spätes Geburtereignis bei geringer Parität protektiv auf die Lebenserwartung aus. Der lebenslaufspezifische Reproduktionserfolg der Frauen hatte demzufolge auch die höchste Effizienz in der Altersgruppe der 70–79-Jährigen und ist im oberen Alter (über 80 Jahre) dann wieder gesunken. Bei den Männern dagegen zeigte sich ein weiterer Anstieg des Reproduktionserfolgs bei zunehmendem Lebensalter. Ebenso war die Überlebenswahrscheinlichkeit der Nachkommen von jenen Elternteilen niedriger, welche bereits im jungen Erwachsenenalter verstorben waren, sie war aber insgesamt höher, sobald die Eltern bis zum einem höheren Lebensalter überlebt hatten.

Erneut zeigte sich, dass die Bedeutsamkeit des Erreichens eines höheren Alters von Seiten der Mütter einen wesentlich stärkeren Einfluss auf das Überleben der Nachkommen hatte als der elterliche Effekt der Väter. Als mögliche Schlussfolgerungen könnte man potentielle Wechselbeziehungen zwischen dem Überleben der Nachkommengeneration und den Eltern mutmaßen, solche konnten aber aufgrund der bisherigen Untersuchungsergebnisse nicht nachgewiesen werden.

Die Selektionseffekte, welche die Fertilität und Mortalität betreffen, könnten auch eine mögliche Erklärung dafür sein, warum die individuelle Fitness in den bürgerlichen Familien seit 1800 gesunken ist, obwohl diese aufgrund ihres sozioökonomischen Status bessere Lebensbedingungen haben müssten. Für beide Bevölkerungsgruppen konnte aber ein Anstieg der Lebenserwartung beobachtet werden, ebenso ein Rückgang der durchschnittlichen Kinderzahl, welche einhergeht mit der sinkenden Zahl von Sterbefällen im Kindesalter.

Korpelainen (2000) führt dabei drei wichtige Erklärungsfaktoren an, welche die individuelle Fitness und den Reproduktionserfolg maßgeblich beeinflussen, und auch in anderen wissenschaftlichen Arbeiten mit ähnlicher Thematik aufgeführt werden. Demzufolge können die Lebenserwartung der Nachkommen, der sozioökonomische Status sowie die Zugehörigkeit eines Geburtsjahrganges die individuelle Gesundheit und das Reproduktionsergebnis beeinflussen.

Insgesamt war der Einfluss dieser Kontrollgröße bei den Frauen stärker ausgeprägt als bei den Männern, welches sich abermals mit den geringeren Reproduktionskosten der männlichen Elternteile erklären lässt.

Ähnliche statistische Zusammenhänge konnten Gundy und Kravdal (2008) in einer Langzeitstudie zu den Fertilitätsbiographien und der Sterblichkeit im mittleren Lebensalter für norwegische Männer und Frauen nachweisen. Die Datenbasis stützt sich dabei auf die Individualinformationen über alle Vitalereignisse des Central Population Register, welches seit der Durchführung des Zensus von 1960 besteht.

Hervorzuheben ist dabei die sehr große Datengrundlage einer landesweiten Vollerhebung von allen Männern und Frauen, die zwischen 1935 und 1958 in Norwegen geboren wurden. Es waren alle Geburtereignisse seit 1965 für jene Personen bekannt, die während der ersten Studienphase zwischen 7 und 30 Jahren alt waren.

Bei den älteren Mitgliedern wurden spezielle Linkage-Verfahren zur durch die Zensen der Nationalstatistiken Norwegens in den 1960er und 1970er Jahre bewirkten Quantifizierung der Eltern-Kind-Beziehungen angewendet, mit einer Ausschöpfungsquote von 81,3% bei den Männern und 77,1% bei den Frauen.

Für die eigentliche empirische Analyse wurden aber nur Personen ab dem Alter von 45 Jahren eingeschlossen, um sicherzustellen, dass insbesondere die Frauen ihre Familienplanung abgeschlossen hatten, und, bei einer festgelegten oberen Altersgrenze von 68 Jahren, die Mitglieder der älteren Geburtsjahrgänge, um die Zeitspanne des Follow-up bis auf Jahr 2003 zu beschränken.

Insgesamt ergibt das eine Fallzahl von 785.317 für die Männer mit 40.071 Sterbefällen und eine absolute Häufigkeit von 744.784 Frauen, die für den Beobachtungszeitraum 23.241 Sterbefälle beigesteuert haben (Grundy und Kravdal, 2008).

Die Ermittlung von Sterblichkeitsunterschieden bei den Frauen und Männern erfolgte durch die Modellierung der Odds-Ratio gemessen an der Anzahl der Kinder, dem Verheiratenstatus und dem Bildungsniveau.

Demnach wirken sich eine höhere Schulbildung und eine Eheschließung protektiv auf das Sterberisiko aus. Bei den Verwitweten sowie bei den geschiedenen bzw. getrennt lebenden Personen ist das Risiko, zu versterben, um ein Vielfaches höher als bei den Singles, wobei dieses wohl mit dem durchschnittlich höheren Lebensalter innerhalb der erstgenannten Personengruppe zu begründen ist. Aufgrund eines Selektionseffekts ist die protektive Wirkung der Ehe auf die Restlebenserwartung bei Männern scheinbar stärker als bei den Frauen.

Dennoch lässt sich mit den Odds für beide Geschlechter zeigen, dass die Sterberisiken für Kinderlose und Personen mit nur einem Kind höher sind als bei Eltern mit zwei Kindern.

Die Sterblichkeitsrisiken sind dann mit steigender Kinderzahl in der Altersgruppe von 45–54 Jahren weiter gesunken, allerdings waren diese Schätzwerte nicht signifikant, was darauf schlussfolgern lässt, dass die Parität zunächst nicht mit der Sterblichkeit in Verbindung gesetzt werden kann. Bei einer entsprechenden Adjustierung der Geburtenniveaus nach Verheiraturungsstatus zeigten sich bei verheirateten Frauen mit drei und mehr Kindern deutlich geringere Sterberisiken als bei Ehefrauen mit nur zwei Kindern.

Bei den verheirateten Männern ergeben sich die stärksten Unterschiede in der Sterblichkeit von jenen, die mehr als fünf Kinder gezeugt hatten, in Relation zu den Männern mit ebenfalls nur zwei Kindern. Ebenso hat das jeweilige Lebensalter der Eltern zum Zeitpunkt des Geburtseignisses einen wichtigen Einfluss auf die Sterblichkeit in der postreproduktiven Phase.

Erneut konnte nachgewiesen werden (vgl. Korpelainen, 2000; Sear, 2007), dass die Mortalitätsrisiken sinken (bei Männern in der Altersgruppe von 55–68 Jahren und bei den Frauen im Alter von 45–54 Jahren), wenn das letzte Geburtseignis in einem Alter von 40 Jahren und höher stattgefunden haben.

1.3.4.2 Negativer Zusammenhang von Fertilität und postreproduktiver Lebenserwartung

Bei der bisherigen Darstellung wurden nur jene Ansätze berücksichtigt, die einen positiven Zusammenhang zwischen der Fertilität und der postreproduktiven Lebenserwartung ermittelt haben. Dabei besteht mindestens eine gleiche Anzahl von wissenschaftlichen Arbeiten, die eine negative Korrelation bzw. einen nichtsignifikanten Gesamteffekt empirisch nachweisen konnten.

Eine zentrale Argumentation zur Wichtigkeit einer umfassenden statistischen Kontrolle von Störgrößen, welche unterschiedliche Selektionseffekte bei der Fertilität und der Sterblichkeit aufzeigen, um somit eine mögliche Unterschätzung des wahren Gesamteffektes zwischen diesen beiden Untersuchungsgrößen zu vermeiden, erfolgte in einer Studie zu der Hollingsworth's Genealogy of the British Peerage (Doblhammer und Oeppen, 2000).

Dieser Datensatz enthält Informationen zur familiären Abstammung von mehr als 4000 Personen (1854 Frauen und 2202 Männern) der britischen Aristokratie über den Zeitraum von 1603 bis 1959. Neben der zeitlichen und kulturellen Veränderung des Stillverhaltens der Mütter wurden auch sozio-ökonomische Unterschiede bezogen auf Einkommen und Bildungsniveau bei der Modellschätzung mit berücksichtigt. Sehr innovativ war auch die Berücksichtigung möglicher Zufallseffekte von unbeobachteter Heterogenität, welche ein wenig kontroverse Schätzergebnisse herbeiführten.

So konnte zunächst für die weibliche Teilpopulation des British Peerage kein statistischer Zusammenhang zwischen der Parität und der Sterblichkeit festgestellt werden.

Durch die Korrektur der unbeobachteten gesundheitsspezifischen Effektgrößen (unbeobachtete Heterogenität), welche auf die Parität und die postreproduktive Sterblichkeit einwirken, konnte eine signifikant positive Assoziation zwischen den beiden Vergleichsgrößen hergestellt werden.

Bei Berücksichtigung der Gesamtzahl aller geborenen Kinder würde sich dann das Sterberisiko um 1,9% ($p=0.24$) bei jedem weiteren Kind erhöhen (Doblhammer und Oeppen, 2000), im restriktiven Modell bei 2 und mehr Kindern steigt dieses Risiko sogar um 3,8%. Bei den Männern konnte zwar ein ähnlicher Trend nachgewiesen werden (das Sterberisiko steigt um 2,7% ($p=0.16$) bei jedem weiteren Kind), dieser war aber statistisch nicht signifikant.

Ähnliche statistische Zusammenhänge konnten in einer Studie zu den Langzeiteffekten von Kinderzahl und Sterblichkeit im vorindustriellen Schweden von Dribe (2004) bestätigt werden.

Für die Studienpopulation wurden insgesamt 2094 Personen eingeschlossen (1063 Männer und 1031 Frauen), diesen konnte in 76% der Fälle ein Sterbeereignis zugeordnet werden, 11% der Personen waren bis zum Studienende im Jahre 1895 noch am Leben und nur 13% sind durch Migration bzw. Zensurierungen ausgeschieden.

Die statistische Analyse des Zusammenhangs zwischen dem Sterberisiko in der Altersgruppe zwischen 50 und 90 Jahren und den Kovariablen wie dem Geburtsjahr, dem sozialen Status im Alter von 50, der Gesamtzahl aller geborenen Kinder, dem Familienstand sowie dem Alter bei der ersten und letzten Geburt erfolgte mit dem Cox-Proportional-Hazard-Modell. Nach Überprüfung der Gültigkeit der Proportionalitätsannahme für den Hazardvergleich durch einen statistischen Test der Schoenfeld-Residuen¹ konnten folgende Ergebnisse ermittelt werden: ein negativer statistischer Zusammenhang zwischen der Kinderzahl und der postreproduktiven Sterblichkeit bei den Frauen und eine statistisch nicht signifikante Assoziation bei den Männern.

Insgesamt war der statistische Effekt bei Frauen mit einer Parität von 4–5 Kindern mit einem 1,52-fach höheren Sterberisiko gegenüber der Referenz von Frauen mit der Parität von 0–1 Kindern am höchsten. Das würde auch bedeuten, dass insbesondere bei den Frauen das Sterberisiko bei steigender Kinderzahl zunehmen würde. Bei den Männern waren die relativen Sterberisiken weniger paritätsabhängig, sodass ein eindeutiger Beweis nicht ermittelt werden konnte.

¹ Gemäß der Annahme von Grambsch and Thernau (1994, 2000) ist die Verteilung der Predictor-Variable identisch mit den Risikosettings. Eine Adjustierung erfolgt durch eine Rechenoperation aus der Varianz-Kovarianz-Matrix der geschätzten Parameter und der Gesamtzahl aller beobachteten Ereignisse des Samples. Eine ausführliche Darstellung der Schoenfeld-Residuals erfolgt im formalen Teil der Dissertation.

Diesen geschlechtsspezifischen Unterschied versucht Dribe (2004) damit zu erklären, dass die negativen Einflüsse durch physiologisch-biologische Belastungen der Schwangerschaft und der Geburt bei Frauen wesentlich stärker sind. Bezogen auf das Alter bei dem ersten Geburtseignis konnte für beide Geschlechter kein statistisch signifikantes Ergebnis festgestellt werden.

Ist die Reproduktion dagegen schon mit relativ jungen Lebensjahren abgeschlossen, so kann dieses zumindest in der weiblichen Teilpopulation zu einer Reduktion der Lebenserwartung in der postreproduktiven Phase ab dem Alter von 50 Jahren führen. Auch hier lässt sich das Argument der Existenz von gesundheitsbedingten Selektionseffekten, welche eine relativ frühzeitige Beendigung der Reproduktionsphase bewirken, anführen.

1.3.4.3 Statistisch nichtsignifikanter Zusammenhang von Fertilität und postreproduktiver Lebenserwartung

Ein Beispiel für statistisch nichtsignifikante Ergebnisse stellt die Arbeit von Helle et al. (2005) dar, welche versucht hat langfristige Korrelationseffekte zwischen der Reproduktion und der Lebenserwartung bei Frauen aus einer historischen Samipopulation aufzuzeigen. Die im Norden Finnlands und Schweden lebenden Sami gelten als eine der wenigen ethnischen Gruppen Europas, die sich durch eine traditionelle Lebensweise auszeichnet, von der Eigenproduktion für den alltäglichen Bedarf lebt und eine relativ hohe Fertilität hat, die keine aktiv betriebene Geburtenbeschränkung kennt.

Die zu untersuchende Datenbasis setzt sich aus 322 Frauen der Sami-Population zusammen, die zwischen 1679 und 1839 in den nördlichsten Regionen Skandinaviens (Utsjoki, Inari und Enontekiö) gelebt haben, nur insgesamt einmal verheiratet waren, ihre Reproduktion weitestgehend abgeschlossen hatten, wobei aber eine genaue Altersangabe bei Erreichen der postreproduktiven Phase nicht genau bestimmt werden konnte.

Das könnte auch eine mögliche Erklärung für die ermittelten Studienergebnisse sein. So hatten Frauen eine höhere Lebenserwartung, wenn sie in einem relativ hohen Lebensalter noch Kinder geboren hatten, genauer reduziert sich das Sterberisiko um jeweils 28% pro Lebensjahr, je länger sie ihre letzten Geburtsergebnisse aufgeschoben haben.

Protektiv auf die Lebenserwartung der Frauen wirken sich auch die verlängerten Überlebenszeiten der Ehepartner aus, da eine verwitwete Frau oftmals bei einer großen Kinderzahl nicht alleine für den Lebensunterhalt aufkommen konnte.

Das Alter bei der ersten Geburt, die Gesamtzahl aller geborenen Kinder sowie die Kohortenzugehörigkeiten hatten keinen signifikanten Einfluss auf die postreproduktive Sterblichkeit. Das bedeutet auch, dass kinderlose Frauen die identischen Sterberisiken aufweisen wie jene, die ein und mehr Kinder geboren hatten. Erstaunlich ist dabei die beobachtbare sehr geringe Müttersterblichkeit (nur 4,5%) unmittelbar nach Geburt, sodass diese kaum einen Einfluss auf die Gesamtsterblichkeit in der Sami-Bevölkerung haben konnte.

Helle et al. (2005) vermuten dabei auch eine Unterschätzung der Langzeiteffekte von Reproduktionskosten bei relativ späten Elternschaften, da es insbesondere nicht unproblematisch ist, eine exakte Quantifizierung von spezifischen Fertilitätsmustern und den Sterberisiken für die Nachkommen von älteren Elternteilen herzuleiten. Möglich sind auch potentielle Selektionsfaktoren, die eine Verlängerung der reproduktiven Lebensphase der Sami-Frauen herbeiführen, sodass zum einem sich Geburten in einer relativ frühen Altersphase, ereignen, zum anderem einige Ereignisse erst im oberen Lebensalter stattfinden.

Eine eventuelle positive Korrelation des Alters der Mütter bei der Geburt des ersten Kindes und dem Alter bei der letzten Geburt setzt eine gleichzeitige Maximierung der Vitalität der Frauen zu einem sehr frühen bzw. sehr späten Zeitpunkt ihrer reproduktiven Phase voraus, was angesichts der zunehmenden körperlichen Beanspruchung nicht sehr realistisch zu sein scheint.

Eine andere Interpretation für eine relativ späte Reproduktionsaktivität könnte sein, dass entstandene Verluste der Fekundität während der frühen Reproduktionsphase kompensiert werden sollen; das setzt aber wiederum voraus, dass sich im Laufe des Lebens eine bestimmte Anzahl von gewünschten Kindern manifestiert.

1.3.5 Zwischenfazit

Mit diesem Zwischenfazit sollen noch einmal alle wesentlichen in diesem Kapitel behandelten Gesichtspunkte in Kurzfassung beschrieben und diskutiert werden.

Zunächst wurden in den ersten beiden Abschnitten einige evolutionsbiologische Ansätze bzw. Theorien sozialer Netzwerke, die mögliche Erklärungsmuster für empirische Zusammenhänge von Fertilität und postreproduktiver Sterblichkeit umschreiben könnten, einander gegenübergestellt. Den evolutionsbiologischen Ansätzen sind die Theorien der Akkumulation von Mutationen (Medawar, 1952), die antagonistische Pleiotrophie (Williams, 1957) und die Disposable-Soma-Theorie (Kirkwood, 1993,1999) zuzuordnen.

Die Theorie der Akkumulationen von Mutationen (Medawar, 1952) geht davon aus, dass die Reproduktionsmöglichkeiten von jenen Individuen geringer ist, die schon in einer sehr frühen Lebensphase durch Krankheiten und Mutationen beeinträchtigt sind.

Treten diese gesundheitlichen Störungen erst relativ spät, also erst in der postreproduktiven Lebensphase auf, dann sind keine nennenswerten Störfaktoren aktiv, welche die Fertilität negativ beeinflussen könnten. Im oberen Lebensalter treten diese Faktoren in massiver Form kumulativ auf und verschlechtern die Gesundheit und die Sterblichkeit nachhaltig.

Die Argumentation von Medawar (1952) beschreibt eine sehr einfache Tautologie eines mit höherem Alter ansteigenden Erkrankungs- bzw. Sterberisikos. Mit der Theorie der antagonistischen Pleiotrophie (Williams, 1957) erfolgte eine wesentlich umfassendere und konkretisierende Darstellung von sich im Lebenslauf verändernden genetischen Komponenten.

So werden für die jeweilige Lebensphase unterschiedliche Gene verantwortlich gemacht, die mehr oder weniger selektiv auf die Gesundheit einwirken. Bestimmte Gene sollen demzufolge die individuelle Fitness bei jungen Erwachsenen zunächst stärken, andere das Leben im Seniorenalter selektiv negativ beeinflussen. Durchaus hat diese theoretische Erklärung ihre Berechtigung, zumal es eine Reihe von wissenschaftlichen Ansätzen gibt (Christensen et al., 2006; Willcox et al., 2006), die eine genetische Erklärung für ein Ansteigen der Lebenserwartung in modernen Gesellschaften feststellen.

Dennoch ist eine einseitige Betrachtung von genetischen Mechanismen für eine vollständige Erklärung einer potenziellen Beziehung zwischen Fertilität und Sterblichkeit nicht ausreichend (Mitteldorf, 2009).

Schließlich beschreibt die Disposable-Soma-Theorie (Kirkwood, 1991,1999) eine evolutionsbiologische Reaktion bzw. eine Anpassung an sich verändernde Umweltbedingungen. In einer Phase von allgemein schlechteren Lebensbedingungen bei einer deutlich geringeren durchschnittlichen Lebenserwartung müssen die Reproduktionsaktivitäten von Organismen aufgrund dieser Verknappung von Lebenszeit in einer viel kürzen Zeitspanne erfolgen.

Sind die äußeren Umweltbedingungen besser und ist die aktiv verbrachte Lebenszeit länger, können die Individuen ihre Reproduktion auf einen späteren Zeitpunkt aufschieben und somit eher ihre körperlichen Aktivitäten auf die Regenerierung der eigenen Physis konzentrieren. Dieses setzt aber zum einen eine stetige Anpassung der Individuen an sich ständig wechselnde Umweltbedingungen voraus, zum anderen erfordert es auch die richtige Entscheidungsfindung in jeder Lebenssituation.

Dabei obliegt es den Individuen, die richtigen physischen Optionen für die jeweiligen Lebensumstände angemessen zu optimieren, was eher spontan erfolgt und daher nicht immer logisch, ökonomisch bzw. altruistisch sein muss. Die evolutionsbiologischen Ansätze beschreiben die rein biologisch-physischen Mechanismen, die eine Erklärung für das Fertilitätsverhalten im Zusammenhang mit Gesundheit und Sterblichkeit liefern sollen.

Soziale Aspekte bzw. die Einflüsse von Netzwerken werden bei ihrer Theoriebildung kaum oder nur ansatzweise berücksichtigt. Insbesondere körperlicher und psychischer Stress, der bei der Beschaffung von lebenswichtigen Ressourcen häufig mit auftritt und die physische Substanz der Individuen schwächt, kann durch soziale Netzwerke kompensiert werden, indem andere Individuen einen Teil ihrer Ressourcen für andere zur Verfügung stellen (Cohen, 2004).

Sind die lebensnotwendigen Ressourcen eher knapp, kann dieser Austausch nicht stattfinden, zumal nicht immer eindeutig geklärt werden kann, ob bzw. inwiefern Individuen für die Bereitstellung von Ressourcen eine Gegenleistung einfordern.

Rosengren et al. (1993) konnten in diesem Zusammenhang zeigen, dass sich die Sterberisiken von Personen verringern, wenn sie auch im oberen Lebensalter nicht in sozialer Isolation lebten, sondern ihre sozialen Kontakte pflegten. Die Wirkung bzw. die Wichtigkeit von sozialen Beziehungen kann dabei sehr unterschiedlich sein. Wie Giles et al. (2005) gezeigt haben, konnte den sozialen Kontakten von älteren Personen zu Freunden und Bekannten ein stärkerer positiver Einfluss auf die Lebenserwartung zugeordnet werden als den Beziehungen zu den eigenen Familienangehörigen.

Im Abschnitt 1.3.3 erfolgte eine Gegenüberstellung von sieben Hypothesen, die sowohl evolutionsbiologische als auch soziale Elemente bei der Untersuchung einer Beziehung von Reproduktion und postreproduktiver Lebenserwartung zulassen.

Smith et al. (2002) konnten dabei einen signifikanten positiven Zusammenhang zwischen dem Alter der Mutter und ihrer Lebenserwartung in Abhängigkeit vom letzten Geburtseignis bei einer niedrigeren Parität ermitteln und damit die Gültigkeit der Theorien des Disposable Soma und der antagonistischen Pleiotrophie bestätigen.

Gagnon et al. (2005) stellten bei einer weiteren empirischen Überprüfung dieser sieben Hypothesen anhand von historischen Individualdaten des Registre de population du Québec und des Programm de recherche en démographie historique (PRDH) fest, dass sich die Berechnungen von Smith et al. (2002) nicht vollständig mit ihren Daten replizieren lassen. Die größten Ergebnisunterschiede ergeben sich gerade bei der Analyse der männlichen Komponente.

Einen positiven Effekt auf die Lebenserwartung von Vätern unter Berücksichtigung eines späten Geburtseignisses bei einer niedrigeren Parität konnte nur in der Utah-Population von Smith et al. 2002 bestimmt werden. Dieser Effekt ist aber nicht mit Daten der frankokanadischen Bevölkerung nachweisbar.

Diese Diskrepanz in den Studienergebnissen findet sich auch in der wissenschaftlichen Diskussion zu dieser Thematik wieder, welches folglich keine eindeutigen Schlussfolgerungen zulässt. Es existiert eine Reihe von wissenschaftlichen Arbeiten, die eine positive Assoziation (Lycett et al., 2000; Mueller et al., 2002; Sear, 2007; Grundy und Kravdal, 2008) von Fertilität gemessen an der Parität und der postreproduktiven Lebenserwartung ermitteln haben.

Ein negativer Zusammenhang konnte empirisch in den Studien von (Doblhammer und Oeppen, 2000; Dribe 2004; Smith et al., 2002; Gagnon et al., 2005, 2008) nachgewiesen werden.

Keine statistisch-signifikanten Zusammenhänge zeigten die Studienergebnisse von Le Bourg (1993) und Helle et al. (2005). Diese teilweise doch sehr voneinander abweichenden Ergebnissen liegen ursächlich in der unterschiedlichen methodischen Herangehensweise in der Messung von Fertilität, in der für die Population verwendeten Datenstruktur sowie in der Beachtung von potentiellen Störungen. Dieses würde auch erklären, warum einige Studienergebnisse (Le Bourg, 1993; Müller et al., 2002; Gagnon, 2008) sich derartig voneinander unterscheiden, obwohl für die statistische Analyse eine identische Datengrundlage verwendet wurde.

Um diese Frage beantworten zu können, gilt es zunächst einmal zu klären, wie die Erklärungsgröße Fertilität gemessen wird. Die Fertilitätsintensitäten können einmal durch die Gesamtzahl aller geborenen Kinder beschrieben werden, welche aber keine Auskunft über den Reproduktionserfolg gibt. Da die meisten Studien auf historischen Daten von Populationen der vorindustriellen Epoche beruhen, sind aufgrund der noch bestehenden sehr hohen Säuglings- und Kindersterblichkeit möglicherweise die Schätzergebnisse verzerrt, sobald die Fertilität nur anhand der Parität festgemacht wird.

Langfristige Effekte der Fertilität auf die postreproduktive Sterblichkeit könnten sich eindeutig besser bestimmen lassen, wenn, wie im Falle der wissenschaftlichen Arbeit von Korpelainen (2000), die Zahl der überlebenden Kinder bis zum 18. Lebensjahr als Indikator für die Fertilitätsintensitäten herangezogen wird.

Die Variation innerhalb Studienergebnisse könnte auch verursacht werden durch die sehr unterschiedliche Herangehensweise bei der Beachtung von potentiellen Störgrößen wie z.B. dem Bildungsniveau und dem sozioökonomischen Status der Eltern, die einen Einfluss sowohl auf die Fertilität als auch auf die Sterblichkeit haben könnten.

Doblhammer und Oeppen (2002) haben in ihren Arbeit anhand von zwei unterschiedlichen Modellrechnungen, einmal mit und einmal ohne Beachtung des Aspektes von unbeobachteter Heterogenität im Hinblick auf Selektionsmechanismen, welche die Gesundheit betreffen, jeweils unterschiedliche Interpretationsmöglichkeiten für den Gesamteffekt eröffnet.

Beide Autoren betonen besonders die Gefahr von bei allen wissenschaftlichen Studien möglichen Fehlinterpretationen, sobald die unterschiedlichen Bevölkerungsstrukturen, die sich nicht immer direkt beobachten lassen, nur unzureichend bei der statistischen Analyse beachtet werden.

Auffällig ist dabei auch, dass viele wissenschaftliche Studien (Lycett et al., 2000; Dribe, 2004; Helle et al., 2005) die männliche Perspektive eines Zusammenwirkens von Fertilität und postreproduktiver Sterblichkeit nur sehr geringfügig oder überhaupt nicht berücksichtigen.

Ein eindeutiger Zusammenhang zwischen der Fertilität und postreproduktiven Sterblichkeit von Vätern lässt sich nur in den Arbeiten von Smith et al. (2002) sowie bei Grundy und Kravdal (2008) ermitteln, bei anderen Ansätzen (Dribe, 2004; Gagnon, 2008) konnten keine signifikanten Sterblichkeitsunterschiede von kinderlosen und Männern mit Kindern nachgewiesen werden.

Begründet wird dieses mit den nachweisbaren unterschiedlichen Reproduktionskosten von Männern und Frauen und der weniger engen körperlichen Verbindung der Väter zu ihren Kindern. Gagnon et al. (2005) argumentieren, dass gerade die männlichen Sterblichkeitsvorteile indirekt mit der relativ großen Kinderzahl der Frauen zusammenhängen. Für die Männer war die protektive Altersvorsorge auch im oberen Lebensalter besser gesichert, sobald mehrere Kinder das Erwachsenenalter erreicht hatten und somit ihre Väter wirtschaftlich mitversorgen konnten.

Offen bleibt weiterhin die Wichtigkeit der biologischen Vaterschaft im Zusammenhang mit der postreproduktiven Sterblichkeit, zumal Kinder gerade in der heutigen Gesellschaft auch aus vorherigen Eheschließungen und nicht-ehelichen Lebensgemeinschaften entstammen könnten bzw. hätten adoptiert werden können.

Gerade die letzte Argumentation verdeutlicht die Relevanz bzw. die Bezugsmöglichkeiten von statistischen Schätzungen auf Basis historischer Populationen für die aktuellen gesellschaftlichen Strukturen.

Welche Rolle spielen dabei die historisch eng manifestierten Klassenunterschiede und die Bedeutung der Ehe für die Geburtenintensitäten?

Die Ehe als Institution hat nach wie vor eine protektive langfristige Wirkung auf die Lebenserwartung. Aus historischer Perspektive konnte eine Ehe nur durch den Tod eines Partners aufgelöst werden, in den heutigen Gesellschaften sind die Ehen weniger stabil.

Beinahe jede dritte Ehe (entspricht 190.900 Ehescheidungen) wurde in Deutschland im Jahr 2008 geschieden (Statistisches Bundesamt, 2008). Hierbei handelt es um eine reine Querschnittsbetrachtung, eine bessere Aussagefähigkeit mit statistischen Informationen aus dem Längsschnitt ist nur eingeschränkt möglich, da diese Zahlen zu den Ehescheidungen für Deutschland nur begrenzt verfügbar sind.

Der Anteil der nichtehelichen Geburten ist in den letzten 20 Jahren deutlich angestiegen. In den neuen Bundesländern haben die Geburten außerhalb der Ehe seit dem Jahr 2000 eine größere Bedeutung als die ehelichen Geburten und entsprechen daher einem relativen Anteil von etwa 60% (Statistisches Bundesamt, 2007).

Auch im früheren Bundesgebiet haben sich die Anteilswerte der nichtehelichen Geburten im Vergleich vom Jahr 1990 mit 10% auf rund 24% im Jahr 2006 mehr als verdoppelt.

Diese Entwicklung zeigt eine deutliche „Entkoppelung von Familiengründung und Eheschließungen in den Lebensverläufen von Frauen und Männern. Hinter dem Phänomen nichtehelicher Elternschaft verbergen sich unterschiedliche ‚neue Familienformen‘, die auch in sozialstruktureller Hinsicht heterogen sind. Entsprechend können die Ursachen der Zunahme der nichtehelichen Elternschaft variieren.“ (Konietzka und Kreienfeld 2005, S. 34). Neben der Vielfältigkeit neuer Familienformen ist auch ein deutlicher Geburtenrückgang in modernen industriellen Gesellschaften beobachtbar.

Daher ist auch eher fraglich, ob ein Vergleich von sehr traditionell lebenden Populationen aus früheren historischen Zeitphasen vorzunehmen ist, bei denen eine deutlich größere Verbindung bzw. Abhängigkeit von der natürlichen Umgebung nachweisbar ist und diese dann sowohl die Fertilität als auch die Sterblichkeit erheblich beeinflussen könnte (Thomas et al., 2000).

Untersuchungspopulationen mit einer sehr hohen Parität, beispielsweise der frankokanadischen Bevölkerung in der Studie von Müller et al. (2002) zeigten dass keine aktiv betriebene Geburtenkontrolle erfolgte und angesichts der hohen Kinderzahlen (78,3% der Mütter hatten bis zu 4 Kinder, 54,8% bis zu 8 Kinder und 21,6% sogar bis zu 12 Kinder) eine Frau nur dann kinderlos blieb, wenn sie nicht verheiratet bzw. physisch nicht in der Lage war Kinder zu gebären.

1.3.6 Zusammenhang von Fertilität und postreproduktiver Sterblichkeit bei Männern mit Spermogramm-Informationen

Bis zum aktuellen Stand der Forschung existieren im Wesentlichen zwei wissenschaftliche Arbeiten, welche einen statistischen Zusammenhang zwischen der männlichen Fertilität und Sterblichkeit anhand von klinischen Spermogrammdaten ermittelt haben.

Die Arbeit von Groos et al. 2006 thematisiert die Untersuchung der Lebenszeit-Mortalität von subfertilen und fertilen Männern.

Sie gilt als Vorlaufstudie für diese wissenschaftliche Arbeit, beschränkt sich aber bei der statistischen Analyse auf nur 601 Untersuchungspersonen, welche in der Zeitspanne von 1892 und 1937 geboren wurden. Die Angaben zu den Spermogramm-Informationen stammen aus der Andrologischen Abteilung des Universitätsklinikums Marburg und wurden in den Untersuchungsjahren von 1949 bis 1985 erhoben. Die Informationen zum Vitalstatus wurden über die Einwohnermeldeämter und eine hessische Zentraldatei einer Krankenkasse erhoben.

Die Ergebnisse der nichtparametrischen Verfahren zeigten dabei nur einen statistisch signifikanten Zusammenhang für die älteren Geburtsjahrgänge (vor dem 25.08.1931 geboren), wobei sich die größten Unterschiede im Survival-Vergleich bei den normozoospermen und oligozoospermen Fällen zeigten. Für die jüngeren Geburtsjahrgänge konnte dieser Zusammenhang nicht ermittelt werden.

Die semiparametrischen Verfahren der Cox-Regression bestimmten ein 1,6-mal höheres Sterberisiko der subfertilen gegenüber der fertilen Vergleichsgruppe und ein doppeltes Risiko der oligozoospermen Männer im Vergleich zu den normozoospermen Patienten.

Die vollparametrischen Verfahren wie Gompertz-Makeham, Weibull und Exponentialverteilung führten zu vergleichbaren Ergebnissen, kleinste Abweichungen in den Parameterschätzungen sind auf die unterschiedlichen Spezifikationsannahmen des Baseline-Risikos zurückzuführen.

Bezüglich der Studienergebnisse von Groos (2005) sind aber einige Gesichtspunkte kritisch anzumerken.

Für die Analyse wurde nur die gesamte Lebenszeit (Lebensspanne zwischen der Geburt und dem Todeszeitpunkt) behandelt, nicht aber die Zeit zwischen der ersten Untersuchung und dem Beobachtungsende (die postreproduktive Phase). Somit wird unterstellt, dass bei allen Personen, bei denen im Laufe des Lebens Fertilitätsstörungen aufgetreten sind, diese schon zum Zeitpunkt der Geburt vorhanden waren.

Die Betrachtung der Einflüsse eines so genannten Kohorteneffektes und der damit verbundenen Heterogenität erfolgte nur unzureichend. Es ist bekannt, dass im direkten Vergleich der kohortenspezifischen Sterblichkeit jüngere Kohorten, bedingt durch den Mortalitätsfortschritt, insgesamt deutlich bessere Überlebenschancen haben als ältere Geburtsjahrgänge.

Im relativen Vergleich der altersspezifischen Sterbewahrscheinlichkeiten für die Kohorten 1892 und 1937 wäre für das Beispiel eines 50-jährigen Mannes, gemessen am relativen Survival, eine Verbesserung der Mortalität um 42 Prozent (Statistisches Bundesamt, 2006) zu erwarten.

Deshalb ist zu vermuten, dass der Sterblichkeitsunterschied von normozoospermen und oligozoospermen Männern für die älteren Geburtsjahrgänge nur deshalb nachweisbar ist, weil die Sterblichkeit bei diesen Personen insgesamt höher ist als in allen anderen nachfolgenden Jahrgängen.

Diese Argumentation setzt aber auch voraus, dass der Geburtsjahrgang ein Prädiktor für die Fertilität bei Männern sei. Problematisch ist dabei, dass, bedingt durch das Follow-up-Studiendesign, sich die kohortenspezifische Mortalität nicht exakt bestimmen lässt.

In diesem Zusammenhang müssen für jene Geburtsjahrgänge, welche vor dem Jahr 1920 geboren wurden, weitere Lebensrisiken wie die zwei Weltkriege, Epidemien wie z.B. die spanische Grippe (1918–1920) sowie die wirtschaftlichen und sozialen Folgen der Depression des Jahres 1929 beachtet werden.

Bei den parametrischen Regressionsverfahren verweist Groos (2005) zwar auf die Schätzung für die p-Werte, die Likelihood-Parameter für die jeweiligen Modelle wurden aber nicht angegeben, sodass unklar bleibt, welche der Spezifikationsannahmen für die Baseline-Hazards zumindest nach der Frequentist-Logik² (Hampel, 1998) am besten das gesamte Modell beschreiben, unabhängig davon, ob dafür ein signifikanter Unterschied zu ermitteln ist.

In einer aktuellen Studie wurde der systematische Zusammenhang zwischen der Spermienqualität und der Lebenserwartung für eine Kohorte von 43277 Männern bestimmt (Jensen et al., 2009).

Dazu wurden von einem Kopenhagener Andrologie-Labor alle männlichen Probanden rekrutiert, von denen im Zeitraum von 1963–2001 eine oder mehrere Spermienproben untersucht wurden.

Ein hoher Informationsgehalt dieser Studie konnte gewährleistet werden, da die aufbereiteten Individualdaten zur männlichen Fertilität mittels einer personalisierten Identifikationsnummer mit anderen Datenquellen wie dem nationalen Sterberegister, der Statistik Dänemark und dem dänischen Krebsregister verknüpft wurden, sodass mit relativ großer Sicherheit eine exakte Auskunft über den Vitalstatus, die erreichte maximale Kinderzahl, die Todesursachen sowie die Begleitiden gegeben werden kann.

² Frequentist-Ansatz: die Anpassung eines statistischen Modells erfolgt durch Maximierung der Likelihood-Funktion.

Hervorzuheben ist dabei auch, dass sich die Studienergebnisse auf eine relativ große Fallzahl, eine sehr lange Follow-up-Periode und auf eine vollständige Erfassungsgenauigkeit durch die Registerdaten beziehen. Bezüglich der Todesursachenstatistik sollten einige Fehlklassifikationen auftreten, diese würden gemäß der Argumentation der Autoren (Jensen et al., 2009) keinen Einfluss auf das Studienergebnis haben. Die Effektgröße Spermienqualität sollte auch in diesem Kontext nicht als reiner Fertilitätsparameter interpretiert werden, da nicht ausgeschlossen werden kann, dass sich einige männliche Patienten bewusst gegen Kinder entscheiden, obwohl keine physischen Fertilitätsstörungen vorliegen.

Insgesamt konnte ein Zusammenhang zwischen der Spermienqualität und der Lebenserwartung bestimmt werden, eine deutliche Reduktion der Sterblichkeit bis zu einem Parameterwert von 40 Mio. Spermien pro Milliliter ist zu erkennen.

Kritisch zu beurteilen ist die Tatsache, dass für die Analyse alle azoospermen Fälle ausgeschlossen wurden, da für diese im Verhältnis zu der infertilen Referenzgruppe eine geringere Sterblichkeit festgestellt wurde.

Die Argumentation der Autoren zielt dabei auf die heterogene Zusammensetzung der azoospermen Untersuchungsgruppe, welche nicht ganz nachvollziehbar ist, da zugleich auch Personen, die sich einer Vasektomie unterzogen hatten bei der statistischen Analyse mitberücksichtigt wurden.

Generell ist die Sterblichkeit innerhalb der Studienpopulation als zu günstig zu bewerten und somit nicht mit der allgemeinen dänischen Bevölkerung vergleichbar.

Die Ursachen der Heterogenität, die dieser Untersuchung zum Thema statistischer Untersuchungen dienen, sind natürlich zunächst in der Pathologie und Pathophysiologie des männlichen Reproduktionssystems zu suchen, vielleicht aber auch in der Steuerung der allgemeinen Lebens- und Alterungsprozesse. Eine parallel zu dieser Studie mit dem gleichen Datensatz erstellte Dissertation von Seydel (2010) fand keinen Zusammenhang zwischen der höheren Mortalität der subfertilen Fälle und einem bestimmten Todesursachenprofil. Folglich ist vermutlich keine statistische Assoziation für ein bestimmtes Morbiditätsmuster erkennbar.

Annahmen über die Ursachen der Mortalitätsunterschiede ebenso wie der ja nur teilweise zu beobachtenden Heterogenität werden zu diesem Zeitpunkt also spekulativ bleiben müssen, um so wichtiger und sinnvoller scheint es für den Autor, einen kurzen Überblick über die normale Struktur und Funktion des männlichen Reproduktionssystems zu geben, damit der Leser die teilweise spekulativ bleibende Diskussion möglicher Ursachen und Erklärungsfaktoren für unbeobachtete Heterogenität kritisch einordnen kann.

Bei der Darstellung der Physiologie der Spermiogenese wird einschlägige Literatur – siehe Krause und Weidner 1998, S. 9; Nieschlag et al. 2009 S. 16 ff.; Silbernagel 2001, S. 306–307; Schröder 2004, S. 8 ff. und Schill et al. 2006 – zitiert.

1.4 Anatomische und physiologische Grundlagen der Spermiogenese

1.4.1 Anatomie des männlichen Genitaltrakts

Anatomisch besteht der männliche Genitaltrakt aus dem Penis mit der Harnröhre, den beiden Hoden (Testes) und den weiteren Anhangsorganen, den Nebenhoden (Epidymis), dem Samenleiter (Ductus deferens), den Bläschen-drüsen (Glandulae vesiculariae), der Vorsteherdrüse (Prostata) sowie den Cowper'schen Drüsen.

Der Hoden mit einem Volumen von ca. 20–25 ml wird von einer Bindege-webskapsel (Tunica albuginea) umgeben, welche an ein Mesothel (Epidorchium) angrenzt. Die radiär verlaufenden Septen (Septula testis) unterteilen das Hodenparenchym in ca. 250–350 Läppchen (Lobuli testis). In den Lobuli verlaufen ca. 500–800 schlaufenförmige Kanälchen, die so ge-nannten Keimtubli, welche mit einem Durchmesser von ca. 180–280 Mikro-metern (Krause und Weidner, 1998) für die Produktion der männlichen Keimzellen, die Spermatogenese, verantwortlich sind. Die Keimtubuli sind über die Rete testis, ein netzwerkartiges Gebilde, mit den Nebenhoden ver-bunden. In einem Zwischenraum verlaufen Blut- und Lymphgefäße sowie Nervenfasern. Ebenso befinden sich dort auch die steroiden Zwischenzellen (Leydig-Zellen), welche das Sexualhormon Testosteron produzieren.

1.4.2 Regulierung der Sexualhormonsynthese

Die Regulierung der Fortpflanzungsmechanismen des Mannes erfolgt über den Hypothalamus, der durch pulsatile Ausschüttung von Gonadotrophin-Releasing-Hormonen (GnRH) eine weitere Synthese bzw. Ausschüttung der gonadotropen Hormone, des follikelstimulierenden Hormons (FSH) und des luteinisierenden Hormons (LH) (Nieschlag et al., 2009), beschleunigt.

Der Hypothalamus wird über das übergeordnete Zentralnervensystem gesteuert und seine hormonale Sekretion unterliegt den Feedback-Reaktionen mit den Hoden. Das luteinisierende Hormon (LH) stimuliert die Leydig-Zwischenzellen bei der Produktion des Steroids Testosteron, welches direkt auf die Reifung der Keimzellen einwirkt (Schräder, 2004).

Das follikelstimulierende Hormon (FSH) reguliert den Ablauf der Keimzellenreifung über die Sertoli-Zellen. Durch die Wechselwirkungen des gonadal gebildeten Testosterons, der im peripheren Gewebe gebildeten Östrogene und durch die in den Sertolizellen gebildeten Proteohormone Aktivin und Inhibin B wird die Sekretion von GnRH, LH und FSH (Nieschlag et al., 2009) gesteuert.

1.4.3 Funktion des Hodens

Der Hoden mit seinen endo- und exokrinen Aufgaben übernimmt nicht nur wichtige Regulierungsfunktionen des Stoffwechsels im männlichen Organismus, sondern in ihm finden auch die Spermatogenese (d.h. die Ausbildung der Keimzellen durch mitotische und meiotische Teilung) sowie die Spermiogenese (d.h. die Differenzierung der runden Spermatiden zu Spermatozoen) statt (Schräder, 2004). In den Hoden wird endokrin das männliche Sexualhormon (Androgen) Testosteron gebildet. Das Testosteron wird durch das Einwirken von LH in den Leydig-Zellen gebildet und ruft im männlichen Körper gonadale und extragonadale Effekte hervor.

Zu den extragonadalen Effekten zählen unter anderem die eiweißanabolen Einflüsse (Silbernagel et al., 2001), welche für die Herausbildung der Muskelmasse und den Knochenbau verantwortlich sind. Der maskuline Behaarungstyp sowie das Aggressions- und Sexualverhalten werden ebenfalls durch das Testosteron reguliert.

Für die beiden wichtigsten Reproduktionsmechanismen Spermato- und Spermiogenese müssen zunächst befruchtungsfähige haploide Spermatozoen in den Hoden gebildet werden.

Die dazu erforderliche meiotische und mitotische Teilung und die daran anschließende Differenzierung aus diploiden Spermatogonien erfolgen in den Tubuli seminiferi contorti.

Der immunologische Schutz einer ungestörten Spermatogenese wird erreicht durch die Formierung der so genannten Blut-Hoden-Schranke in den Sertoli-Zellen. In den Keimtubuli läuft die Spermatogenese innerhalb eines Zyklus von 16 Tagen, wobei die Entwicklung eines vollständigen Spermiums ca. 4,6 Zyklen (entspricht 74 Tagen) benötigt, wobei aber noch während der Teilungsphase schon etwa die Hälfte der Spermienzellen abstirbt (Nieschlag et al., 2009; Schill et al. 2006).

Nach Durchwanderung des Nebenhodens, welche etwa 12 Tage andauert, erhalten die Spermatozoen ihre vollständige Motilität und ihre Befruchtungsfähigkeit (Schill et al., 2006). Die befruchtungsfähigen Spermien werden größtenteils im distalen Abschnitt gespeichert und bei der Ejakulation über die ableitenden Samenwege mit den Sekreten der Samenbläschen, der Prostata und weiteren akzessorischen Drüsen im Ejakulat abgegeben.

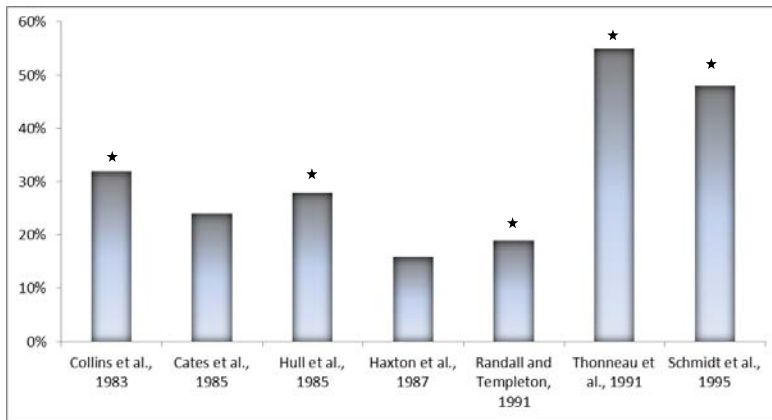
1.4.4 Männliche Infertilität

Von ungewollter Kinderlosigkeit sind in Deutschland beinahe 15% aller Paare im re-produktiven Alter betroffen (Bruckert 1991; Juul et al., 1999; Turek, 2008). Nach Schätzungen der WHO sind weltweit zwischen 60 und 80 Millionen Paaren von ungewollter Kinderlosigkeit betroffen (Fathalla, 1992).

In einer Vielzahl von Veröffentlichungen zu diesem Thema wird allgemein davon ausgegangen, dass der männliche Faktor zu 20% (Nallela und Rakesh, 2006) in allen Fällen allein verantwortlich für die ungewollte Kinderlosigkeit sei und in 30–40% der Fälle einen maßgeblichen Anteil zur Nichterfüllung des Kinderwunsches der Paare beitrage.

Doch die Ergebnisse vorangegangener Studien (Irvine, 1998) beweisen, dass diese Zahlen nicht als allgemeingültig verstanden werden können. Die Werte des männlichen Anteils variieren zwischen knapp 60% (Thonneau et al., 1991) und unter 20% (Haxton et al., 1995).

Abbildung 1: Bedeutung des männlichen Faktors für die Fertilitätsbestimmung



* ein männlicher Faktor als häufigste Einzelursache (Irvine, 1998)

Die Ursachen männlicher Infertilität können nach prätestikulärem, testikulärem und postestikulärem Ursprung aufgegliedert werden.

Tabelle 1: Ursachen der männlichen Unfruchtbarkeit (Dorn, 2007)

Idiopathische Subfertilität	1,7%
Varikozele	16,6%
Infektionen	9,0%
Hypogonadismus	8,9%
Lageanomalie des Testes	8,5%
Störungen der Samendeposition	5,8%
Allgemeiner Erkrankungen	5,0%
Immunologische Faktoren	4,2%
Hodentumore	2,3%
Obstruktionen	1,5%
Rest	6,6%

Nicht berücksichtigt werden Fertilitätseinschränkungen, die unmittelbar nach jeder schweren Infektionskrankheit vorliegen. Ebenso zu beachten ist der Anteil idiopathischer Infertilität an den ungeklärten Ursachen für Infertilität, was bei 30% aller infertilen Männer vorkommt (Nieschlag, 1993).

Somit sind neben idiopathischen Ursachen Infektionen, Varikozelen, Hoden-hochstand und erektile Dysfunktionen mit nahezu 80% aller Fälle die häufigsten Diagnosen bei der männlichen Infertilität.

1.5 Einige andrologische Erkrankungen

1.5.1 Nebenhodenentzündung (Epididymitis)

Eine sehr häufige Erkrankung ist die meist einseitig auftretende Entzündung der Nebenhoden, welche durch Rötung, Schwellung und durch intensive Schmerzen im Hodenbereich gekennzeichnet ist. Ausgelöst wird die Epididymitis durch Keime wie Colibakterien, Enterokokken, Proteuskeime etc. Insbesondere bei jüngeren Patienten mit häufig wechselnden Sexualpartnern sind eher Gonokokken (Trippererreger) oder Chlamydien ursächlich verantwortlich.

Bei unkomplizierten bakteriellen Nebenhodenentzündungen ist eine hochdosierte resistenzgerechte antibiotische Therapie sehr erfolgreich, bei der in den meisten Fällen eine Heilung nach ca. 2–3 Wochen eintritt.

In Fällen, in denen eine Abszedierung (Einschmelzung von Nebenhodengewebe) eintritt, müssen die Nebenhoden operativ entfernt werden (Epididymektomie). Die bakteriellen Epididymitiden verlaufen in 15–20% der Fälle chronisch (Luzzi und O'Brien, 2001, MedHost, 2008). Nach Ausheilung einer Nebenhodenentzündung kann der Nebenhodenkanal verschlossen bleiben und somit die männliche Fertilität beeinträchtigen.

1.5.2 Hodenentzündung

Die Mumpsorchitis ist die häufigste Form der Hodenentzündung und tritt im Erwachsenenalter als Begleit leiden einer Speicheldrüsenentzündung in der Wange (Parotitis) auf. In 30% der Fälle tritt Mumpsorchitis beidseitig auf, es folgt darauf eine irreversible Schädigung der Hoden (testikuläre Atrophie) mit verbleibender Sterilität (Hautmann und Hulan, 2006; Albrecht, 2008).

Neben der Mumpsorchitis können auch bakterielle Allgemeinerkrankungen wie Typhus, Tuberkulose und Syphilis auftreten. In den meisten Krankheitsfällen kann zwar die Ursache antibiotisch therapiert werden, der Funktionsverlust des Hodens ist aber irreversibel.

1.5.3 Urogenitaltuberkulose

Die Prävalenz dieser Krankheitsform ist zum gegenwärtigen Zeitpunkt doch eher gering, dennoch kann gerade bei der anschließenden Untersuchung von Patienten bis einschließlich Geburtsjahr 1941 nicht ausgeschlossen werden, dass jene Personen zu einem früheren Zeitpunkt der Tuberkulose ausgesetzt waren. Sobald die Latenzzeit der reaktivierten Tuberkulose relativ lange dauert, kann einer Streuung von Tuberkelbazillen in den Blutwegen diagnostiziert werden, welche dann „tuberkulöse“ Einlagerungen in den Nieren hervorrufen kann.

Diese Infektion kann je nach Abwehrlage recht schnell abheilen und der Patient bleibt über Jahre beschwerdefrei (Wise und Marella, 2003; DocCheck, 2008). Wenn aber zu einem späteren Zeitpunkt eine erneute tuberkulöse Virulenz in den ableitenden Harnsystemen auftritt, führt diese oftmals zu schweren Läsionen an den Nieren und der Blase, möglicherweise überträgt sich dieses dann auf die Anhangsorgane des Harntraktes wie Prostata, Samenblase oder Nebenhoden. Die Urogenitaltuberkulose kann nur primär antibiotisch behandelt werden.

1.5.4 Varikozele

Als eine krampfaderähnliche Erweiterung der Hodenvenen im Samenstrang wird die Varikozele bei 10–25% aller Männer diagnostiziert und gilt somit als eine der häufigsten Ursachen für männliche Infertilität (Canales et al. 2005; TK-online, 2008). Unterschiedliche Abflussverhältnisse in beiden Hoden (die Hodenvene des rechten Hodens mündet spitzwinkelig in die untere Hohlvene, die linke Hodenvene dagegen rechtwinklig in die linke Nierenvene) bewirken, dass bei klinischen Untersuchungen in 80–90% die Varikozele nur auf der linken Seite, in jeweils 5–10% auf der rechten oder aber auf beiden Seiten auftritt.

Es werden drei Schweregrade (I-III) der Varikozele unterschieden, wobei die dritte Stufe mit einer Verkleinerung des Hodens (Hodenathrophie) verbunden ist.

Dabei gilt: Je ausgeprägter eine Varikozele ist, desto höher ist die Wahrscheinlichkeit einer Infertilität; es bedeutet auch, dass subklinische Varikoze-len normalerweise nicht zu klinisch relevanten Einschränkungen der Fertilität führen (Nagle, 2004).

In der wissenschaftlichen Literatur aufgeführte Argumentationen verweisen auf die mangelnde Reliabilität von parallel durchgeführten prospektiven Studien und der damit verbundenen Abschätzung der Heilungschancen bei klinischer Varikozele und einer möglichen Verbesserung der Fertilität. Der so genannte Cochrane Review (Ficarra et al., 2006) nimmt eine Bewertung der Fachliteratur vor und zeigt, dass kaum Anhaltspunkte für die Effizienz der Behandlungsmöglichkeiten und eine anschließende Verbesserung der Fertilität existieren. Dagegen spricht, dass bei einer frühzeitig angesetzten adäquaten Therapieform für eine weniger stark ausgeprägte Varikozele (Grad 1) die Erfolgchancen einer vollständigen Wiederherstellung der Fertilität bereits nach einem Jahr sehr hoch sind (Schill et al., 2006).

1.5.5 Hypospadie

Bei der Hypospadie handelt es sich um eine angeborene Fehlbildung der abnormen ventralen Mündung der Harnröhre, verbunden mit einer ebenfalls häufigen ventralen Verkrümmung des Penis (Baskin, 2000). Die Ausbildung einer Hypospadie ist bei fehlender oder unzureichender Androgensynthese bei einem Enzymdefekt (10–70%, je nach Schweregrad) wahrscheinlicher.

Als wichtige Risikofaktoren sind das Alter der Mutter, ein niedrigeres Geburtsgewicht sowie eine erhöhte Exposition gegenüber Substanzen mit starker östrogenen Wirkung, wie Pestiziden, PVC-haltigen Kunststoffen, aber auch natürlich gebildetem Östrogen in der Nahrung (siehe Diskussion). Häufige Begleiterleiden sind Kryptochismus, Leistenhernie sowie Fehlbildungen des Herzens. Mit nur etwa 8 von 1000 Fällen gilt die Hypospadie als eine eher seltene andrologische Erkrankung (Springer und Horcher, 2009).

1.5.6 Kryptochismus

Hierbei handelt es sich um eine abnormale Lage des Hodens auf dem Weg des natürlichen Herunterwanderns während der embryonalen Entwicklung. Dabei können fünf verschiedene Lokalisationen der Hoden unterschieden werden (Hutson und Hasthorpe, 2005; Kolon et al. 2004; Virtanen und Toppari, 2008):

1. Bauchhoden: Die Testes befinden sich direkt unterhalb der Bauchdecke, innerhalb des inneren Leistenringes.
2. Leistenhoden: Die Testes sind zwischen inneren und äußeren Leistenring angeordnet.

3. Gleithoden: Die Testes sind ähnlich angeordnet wie beim Leistenhoden, können allerdings bis in den Leistenansatz bewegt werden, anschließend aber wieder in die ursprüngliche Leistenposition gebracht werden.
4. Ektope Hoden: Die Testes liegen abseits der physiologischen Deszensusstrecke.
5. Pendelhoden: Die Testes können zeitweise im Leistenkanal liegen, bei Entspannung bewegen sie sich wieder in die ursprüngliche Lage zurück.

Der unbehandelte Kryptochismus ist unter anderem erkennbar durch: eine geringe Anzahl an Leydig-Zellen, eine strukturelle Veränderung der Sertoli-Zellen, Fehlentwicklungen in den primären Spermatozoen und eine verminderte Keimzellzahl.

Insbesondere bei Frühgeborenen (9–30%) ist der Kryptochismus häufiger präsent. In ca. 20% der Fälle besteht eine Monorchie (einseitiges Fehlen des Hodens) (Virtanen und Toppari, 2008). In Abhängigkeit von der Hodenlokalisierung entwickeln sich die Beeinträchtigungen bei der Keimzellenbildung unterschiedlich intensiv.

Je höher der Kryptochismus lokalisiert ist, desto wahrscheinlicher sind auch Funktionsstörungen der Nebenhoden.

Unklar bleibt aber, inwiefern eine Fehlstellung der Hoden sich auf die Entstehung eines Seminoms auswirkt.

Es wird aber vermutet, dass die Höhe der Hodenlokalisation entscheidend für das Malignitätsrisiko (Meißner et al., 2007) sei.

1.5.7 Hodenkrebs

Beim Hodenkrebs handelt es sich um bösartige Wucherungen des Gewebes, die von den Vorläuferzellen der Spermien ausgehen (Interdisziplinäre Arbeitsgruppe Hodentumore, 2004). Der Befall des gesunden Hodengewebes sowie eine Ausbreitung auf andere Organe kann ohne Behandlung nicht verhindert werden. In der Altersgruppe der 20–40-jährigen Männer tritt diese Erkrankung am häufigsten auf. Jährlich erkranken etwa 7 von 100.000 Männern (Interdisziplinäre Arbeitsgruppe Hodentumore, 2004).

Die Ursachen für die Entstehung für Hodentumore konnten noch nicht eindeutig ermittelt werden. Bekannt ist aber, dass ein diagnostizierter Leisten- oder Pendelhoden bei Männern das Tumorrisiko erhöht. Da die Tumore sich häufig erst nach der Pubertät bilden, wird zur Früherkennung eine Orchidopexie empfohlen, welche, wenn sie noch vor dem 10. Lebensjahr durchgeführt wurde, durchaus sehr effektiv ist.

Die Mehrzahl aller Tumorerkrankungen betrifft in der Regel nur einen Hoden (95%), welcher aber präventiv entfernt werden muss (Albers et al., 2006). Die Sexualität und die Reproduktionsfähigkeit sind dabei meist nicht beeinträchtigt, weil ein funktionsfähiger Hoden genügt, um eine ausreichende Menge an Testosteron und Samenzellen zu produzieren.

Allerdings ist durchaus möglich, dass bei einer Erkrankung der zweite Hoden nur noch eingeschränkt funktioniert und daher die Testosteronkonzentration im Laufe des Lebens schneller abnimmt (Deutsche Krebshilfe, 2009). Müssen dagegen beide Hoden entfernt werden, können keine Spermien mehr gebildet werden, sodass die Patienten dann zeugungsunfähig sind.

Neben den operativen Eingriffen kann auch eine Bestrahlung der Tumorregion im Rahmen einer Chemotherapie³ erfolgen, jedoch muss dabei mit einer Beeinträchtigung der Samenproduktion im noch vorhandenen Hoden gerechnet werden. Welches Ausmaß diese Beeinträchtigungen haben, hängt von der Intensivität der Behandlung ab und erfordert daher einer Einzelfallbehandlung (Deutsche Krebshilfe, 2009).

³ Die Chemotherapie soll das Zellwachstum blockieren sowie in die Zellteilung eingreifen.

1.6 Analyse der Spermiogramm-Parameter

Vor einer Vereinheitlichung gemeinsamer Laborrichtlinien für eine standardisierte Spermiogramm-Untersuchung durch die WHO (1999) waren Vergleiche der Analyseergebnisse von verschiedenen andrologischen Zentren sowohl national als auch international nur sehr schwer möglich. Im Rahmen der Spermiogrammanalyse werden daher einheitlich die drei Standardparameter Spermiedichte, -motilität und -morphologie (Krause, 1993a) bestimmt.

Die Spermiedichte gemessen an der Anzahl der Spermien pro Volumeneinheit Ejakulat wird ermittelt durch die Zählkammermethode (z.B. Neubauer-Zählkammer), indem die Spermatozoen nach einer Verdünnung mit dreiprozentiger Kochsalzlösung im Verhältnis 1:20 immobilisiert werden (Krause und Weidner, 1999).

Die Beurteilung der Beweglichkeit (Motilität) der Spermien erfolgt im Nativpräparat nach drei Bewertungskriterien:

1. Größe des Anteils der beweglichen Spermatozoen an der Gesamtpopulation,
2. verschiedene Beweglichkeitstypen,
3. Höhe der Geschwindigkeit der Spermatozoen.

Die WHO (1999) unterscheidet dabei vier verschiedene Bewegungsstadien:

a) schnell progressiv beweglich, b) langsam progressiv beweglich, c) nicht progressiv beweglich, d) vollständig unbeweglich.

Die Einschätzungen der Motilitätseigenschaften für die Spermatozoen erfolgen überwiegend subjektiv und können je nach Erfahrung des Untersuchungspersonals voneinander abweichen.

Bei der Morphologiebewertung werden die Spermien eingefärbt und auf Defekte an den Köpfen, Hälsen und Mittelstücken untersucht, sodass schließlich der Anteil von morphologisch normalen und anormalen Spermatozoen bestimmt werden kann.

1.6.1 Einfluss des Alters auf die Spermienkonzentration und -qualität

In einer australischen Studie zur Spermienqualität älterer Männer (Ng et al., 2004) wurden 55 infertilen, aber gesunden Männern im Alter zwischen 52 und 79 Jahren im Rahmen eines Screening-Projektes für Prostata-Krebs Spermienproben entnommen.

Wegen der zu geringen Fallzahlen ist eine Extrapolation der Ergebnisse auf die männliche Allgemeinbevölkerung nicht möglich. Die jüngere Kontrollgruppe setzte sich aus 409 freiwilligen Spendern mit einem Alter von < 52 Jahren zusammen, die an einem Fertilisationsprogramm zur künstlichen Befruchtung teilnahmen und aus dem gleichen andrologischen Labor wie die älteren Screening-Patienten rekrutiert wurden. Eine Selektion innerhalb der jüngeren Vergleichsgruppe ist daher nicht auszuschließen, weil die Studienteilnahme freiwillig war und daher vermutlich häufiger jene Personen teilnahmen, die ihren Fertilitätsstatus bereits vor dem Studienbeginn kannten.

Dabei zeigte sich, dass die Spermien-volumen jüngerer Männer mit 3,2 ml gegenüber 1,8 ml ($p < 0,0001$) der älteren männlichen Patienten fast doppelt so hoch sind. Auch in den Spermien-gesamtzahlen unterscheiden sich die beiden Vergleichsgruppen.

Mit 74 Millionen Spermien pro Ejakulat weisen ältere Männer deutlich geringere Werte auf als jüngere Probanden (206 Millionen Spermien/Ejakulat, $p < 0,0001$).

Nichtsignifikante Unterschiede in der Spermien-dichte mit durchschnittlich 64 Millionen und 73 Millionen Spermien pro ml ($p = 0,12$) konnten für beide Gruppen festgestellt werden.

Ältere Männer weisen häufiger abnormale Spermien-morphologien auf, mit sinkenden Anteilen von normalen Formen (durchschnittlich 14 % gegenüber 25%, $P < 0,0001$), einer reduzierten Beweglichkeit (durchschnittlich 51 % gegenüber 80%), einer sinkenden Anzahl von cytoplasmischen Droplets (Median 1 gegenüber 0) und Schwanzabnormalien der Spermien (30% gegenüber 17%, $p < 0,0001$).

Unterschiedliche Anteile der Abnormalien der Kopf- und Mittelstücke konnten zwischen den beiden Vergleichsgruppen nicht festgestellt werden.

1.6.2 Statistische Abgrenzung der Fertilitätsklassifikationen

Die menschliche Spermatogenese wird durch einige scheinbar widersprüchliche Phänomene bestimmt. Beispielsweise weisen einige fertile Männer in ihren Spermogrammen eine Vielzahl von Abnormalien auf, die aber in keiner Weise ihre Fertilität beeinträchtigen.

Werden in den Spermogrammen keine offensichtlichen Abnormalien nachgewiesen, kann ebenfalls nicht von einem uneingeschränkten Fertilitätspotential ausgegangen werden. Folglich wird dann nur in 40% der Fälle eine richtige Diagnose von männlicher Unfruchtbarkeit (Nallella und Rakesh, 2006) vorgenommen.

Die WHO hatte seit 1980 weltweit gültige Standards der Spermienanalyse definiert, diese wurden in unregelmäßigen Abständen (1987, 1992 und 1999) aktualisiert.

Die meisten andrologischen Labore haben diese Standards übernommen, dennoch werden die WHO-Referenzwerte von einigen Klinikern als zu stringent bewertet. Im Jahre 2010 wurden neue Richtlinien für die Beurteilung eines gesunden Spermogramms festgelegt (WHO, 2010). Es gelten nun folgende Referenzwerte.

Tabelle 2: Untere Referenzwerte für die Spermienparameter WHO (2010) ⁴

Ejakulatvolumen	1,5 ml
Gesamtmotilität (progressiv + nichtprogressiv)	40%
Spermiengesamtzahl	39×10^6
Spermienkonzentration	15×10^6
Spermienmorphologie (normale Formen)	4%

In den neuen WHO-Richtlinien (2010) wird ebenfalls darauf hingewiesen, dass im Falle eines Unterschreitens der definierten unteren Referenzwerte bei einer Spermioogramm-Untersuchung nicht unbedingt Infertilität diagnostiziert werden muss.

Es wird daher auch empfohlen, noch andere klinische Informationen für die Beurteilung der Spermienparameter heranzuziehen. Da durchaus auch regionale Unterschiede bezüglich der Spermienqualität auftreten könnten, wird den Laboren ebenfalls empfohlen, eigene angemessene Referenzwerte festzulegen.

Guzick et al. (2001) ermittelten in einer Studie als oberen Grenzwert für Subfertilität eine Spermienkonzentration von 13,5 Mio/Milliliter, einen Anteil motiler Spermien von weniger als 32 Prozent und einen Spermienanteil von weniger als 9 Prozent ohne Abnormalien.

⁴ Laboratory manual for the examination and processing of human semen

Der unterste Grenzwert für Fertilität wurde bei einer Spermienkonzentration von mehr als 48 Mio/Milliliter, mit mehr als 63 Prozent motiler Spermien und mehr als 12 Prozent ohne Abnormalien, operationalisiert. Sollten die Spermienprogrammen-Parameter eines Patienten innerhalb dieses beschriebenen Intervalls liegen, dann besteht nach Auffassung der Wissenschaftler zumindest ein Verdacht für eine Infertilitätsdiagnose.

Eine exakte Abgrenzung von fertilen und subfertilen Fällen erweist sich somit als schwierig. Es kann somit nicht ausgeschlossen werden, dass Männer trotz normaler Spermien-Parameter eine eingeschränkte Fertilität aufweisen und umgekehrt.

Tabelle 3: Statistische Abgrenzung von Fertilität und Subfertilität (Guzick et al., 2001)

Sperm characteristics	FF (n=56)	NHD (n=91)	Infertility evaluation (n=406)	MFI patients (n=166)	PF vs. NHD	PF vs. infertility evaluation	Infertility evaluation vs. NHD	MFI vs. NHD	MFI vs. infertility evaluation	MFI vs. PF
Concentration (10 Mio/mL)*	69.9 (48.3, 120.0)	65.5 (40.7, 95.7)	32.9 (5.0, 75.4)	21.3 (7.95, 44.21)	NS	P<.001	P<.001	P<.001	P<.05	P<.001
Motility (%)	72.5 +/- 16.6	72.9 +/- 15.7	49.3 +/- 30.4	37.0 +/- 20.6	NS	P<.001	P<.001	P<.001	P<.001	P<.001
Morphology (%)										
WHO	37.7 +/- 16.1	35.7 +/- 12.8	17.8 +/- 12.8	29.6 +/- 11.5	NS	P<.001	P<.001	P<.001	P<.01	P<.01
Kruger's strict criteria	13.8 +/- 7.1	10.9 +/- 5.06	5.65 +/- 4.6	8.1 +/- 4.5	NS	P<.05	P<.001	P<.001	P<.01	P<.01

proved fertility (PF), normal healthy donors (NHD), patients undergoing infertility evaluation and patients with male factor infertility (MFI)

Das Spermogramm hat als routinediagnostisches Messinstrument eine zentrale Bedeutung für die Abklärung der männlichen Sterilität.

Es soll dabei zum einen zur Abschätzung des natürlichen Reproduktionspotentials dienen, soll aber zum anderen auch Aufschluss über die Erfolgschancen einer assistierten, reproduktionsmedizinisch herbeigeführten Schwangerschaft liefern. Die Aussagefähigkeit eines solchen Spermiogramms über die Fertilität eines Mannes sei aber nur im Falle der völligen Azoospermie (Runnebaum und Rabe, 1994) gegeben.

Denn im Falle einer Hyperfertilität der Partnerin (die Frau hat besonders gute Fertilitätseigenschaften) ist es nicht ausgeschlossen, dass bei einer diagnostizierten Beeinträchtigung der männlichen Fertilität durchaus eine Schwangerschaft erfolgen kann. Die Vorhersagewahrscheinlichkeit für das Eintreten einer eventuellen Schwangerschaft beträgt nur etwa 23% (Krause, 1993), sodass letztendlich nur eine spontan eintretende Schwangerschaft Auskunft über die tatsächliche Fertilität gibt.

Dennoch ist die klassische Untersuchung des Ejakulats relevant für die Abstimmung von Therapiemaßnahmen zur Behandlung von Infertilität in der Partnerschaft.

Um eine Vergleichbarkeit der Messergebnisse zu gewährleisten und zur Standardisierung bzw. Vereinfachung der Untersuchungsvorgänge empfiehlt sich dabei die Orientierung an den Richtlinien zur Untersuchung des menschlichen Ejakulats der WHO aus dem Jahr 2010.

2. Grundbegriffe der Survivalanalyse

2.1 Definition von Lebensdauer

Die Lebensdauer eines jeden Individuums kann vereinfacht durch eine nicht-negative Zufallsvariable T dargestellt werden, welche mit dem Ereignis Tod endet. Aufgrund ihrer sehr vielseitigen Anwendungsmöglichkeiten muss die Survivalanalyse nicht nur auf die Untersuchung des Ereignisses Tod beschränkt werden. Gerade wegen ihrer methodischen Variabilität kann sie als Ereignisanalyse auch bei anderen demographischen Ereignissen wie dem Auszug aus dem Elternhaus, dem Eintreten der ersten Geburt oder auch in Hinblick auf das zeitliche Eintreten der ersten Eheschließung Anwendung finden. Ebenso lassen sich die Methoden der Survival- oder Ereignisanalyse auch auf andere wissenschaftliche Disziplinen wie zum Beispiel die Medizin, Ingenieur- und Wirtschaftswissenschaften übertragen. In der Medizin wird die Survivalanalyse bei Fragestellungen angewendet, die sich unter anderem auf die Untersuchung von Übertragungswahrscheinlichkeiten schwerer infektiöser Krankheiten oder das Auftreten von postoperativen Komplikationen beziehen. In den Ingenieurwissenschaften wird sie bei der Untersuchung von Halbwertszeiten wichtiger Verschleißteile im Fahrzeugbau genutzt. In den Wirtschaftswissenschaften ist die Ereignisanalyse zum Beispiel bei der Untersuchung von Übergangszeiten aus der Arbeitslosigkeit zurück in den ersten Arbeitsmarkt relevant.

Bei der folgenden Darstellung der Methoden zur Survivalanalyse wird sich auf die Notation von Cleves et al. (2008), Duchateau und Janssen 2008, Hosmer et al. 2008, Klein und Moesberger 2005 sowie die Habilitationsschrift von Wienke (2006) bezogen. Bei betreffenden Stellen werden die Originalquellen mit der entsprechenden Seitenzahl angegeben.

Ausgangspunkt der Survival- bzw. Ereignisanalyse ist unabhängig von der jeweiligen Fragestellung fast immer die kumulative Verteilungsfunktion der Lebensdauer T .

$$F(t) = P(T \leq t) = \int_0^t f(u) du$$

Die Überlebenszeiten auf der Individualebene lassen sich durch die folgende Survivalfunktion (vgl. Duchateau und Jansen 2008, S. 18, sowie Wienke 2006, S. 5–6) beschreiben:

$$S(t) = 1 - F(t) = P(T > t) \quad (2.1)$$

Die Überlebenswahrscheinlichkeiten von Individuen verbleiben zu keinem Zeitpunkt der Lebensspanne bei einem konstanten Niveau, sie unterliegen zu jeder Zeit dynamischen Veränderungen.

Die Hazardfunktion $\lambda(t)$ (auch als Sterberate, Inzidenzrate, Überlebenskurve oder Force of Mortality bezeichnet) charakterisiert die Veränderung des individuellen Sterberisikos über die Zeit und das Alter.

Genauer sie gibt die momentane Neigung eines Individuums an, innerhalb des nächsten Zeitintervalls zu versterben, unter der Bedingung, dass dieses Individuum bis zu diesem Zeitpunkt noch am Leben war.

Die Hazardrate (vgl. Duchateau und Jansen 2008, S. 18, sowie Wienke 2006, S. 5–6) wird folglich definiert als:

$$\lambda(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} = \frac{f(t)}{S(t)} \quad (2.2)$$

Ist die Schätzung von individuellen Hazardraten nur unter Berücksichtigung komplexer Annahmen über die unbekannte Sterblichkeitsneigung innerhalb eines Zeitintervalls möglich,

kann zur Verringerung möglicher Verzerrungen auf die kumulative (oder integrierte) Hazardrate $\Lambda(t)$ zurückgegriffen werden.

$$\Lambda(t) = \int_0^t \lambda(u) du \quad (2.3)$$

Die Survivalfunktion ergibt sich durch die Umformung durch die Terminologie der integrierten Hazardrate⁵ (vgl. Duchateau und Jansen 2008, S. 18–19) und einer anschließenden Exponierung von $\Lambda(t)$ dann wie folgt:

$$S(t) = \exp\left(-\int_0^t \lambda(u) du\right) = \exp(-\Lambda(t)) \quad (2.4)$$

⁵ Die Umformung erfolgt gemäß Duchateau und Jansen, 2008, mit $\lambda(t) = -\frac{d}{dt} \log(S(t))$

Wie bereits erwähnt, ist die Anwendung der Hazardfunktion immer dann vorteilhaft, wenn sich die momentane Sterbeneigung auf der Individualebene über die Zeit verändert.

Kenntnisse über die Musterverläufe der Hazardfunktionen sind wichtig für Quantifizierung und Auswahl des geeigneten Modells, insbesondere bei den parametrischen Schätzverfahren.

Die Hazardfunktionen können verschiedene Formen annehmen: ansteigend, sinkend oder u-förmig sein. Bei demographischen Fragestellungen erhalten die Hazards der Mortalität einen u-förmigen Charakter, gekennzeichnet durch eine erhöhte Perinatalsterblichkeit, hervorgerufen durch mögliche Komplikationen bei der Geburt sowie schwere Erkrankungen in der frühkindlichen Phase, folglich einen minimalen Anstieg der Sterblichkeit, welche dann bis etwa zum 30. Lebensjahr auf niedrigem Niveau konstant bleibt und danach mit zunehmendem Alter wieder ansteigt (Wienke, 2006). In Populationen mit einem relativ hohen Anteil hoch betagter Personen lassen sich ansteigende Hazards anwenden, weil die Sterberisiken mit steigendem Alter schnell ansteigen.

Dagegen sind Modelle mit zunächst sinkenden Hazards in der Demographie bzw. Epidemiologie weniger gebräuchlich, bieten aber in den Ingenieurwissenschaften eine sinnvolle Abbildung der Ausfallwahrscheinlichkeiten von elektrischen Geräten bei relativ langen Laufzeiten.

2.2 Nicht-parametrische Verfahren

Mit diesen Verfahren erfolgt eine direkte Schätzung der empirischen Verteilungsfunktion, ohne die Aufstellung von Annahmen in Form von statistischen Verteilungen über den Sterblichkeitsverlauf.

Wesentliche Vorteile sind dabei die exakte Darstellung der natürlichen Sterbeanordnungen sowie ihre Anpassungstreue an jede statistische Verteilung ohne den Einsatz von komplexen Annahmen. Dagegen können bei nichtparametrischen Verfahren nur bei einem ausreichend großen Datenumfang valide Schätzergebnisse für die Hazard- und Survivalfunktion realisiert werden.

Gerade bei sehr kleinen Fallzahlen erfordern die Schätzergebnisse der diskreten Punktmaße z.B. des Kaplan-Meier-Schätzers den Einsatz von Glättungsprozeduren wie z.B. des Kernel-Smoother. Dieses Verfahren ist immer dann sinnvoll anzuwenden, wenn die empirische Verteilungsfunktion für das Modell unbekannt ist. Dabei wird die maximale Glättung dieser unbekannten Funktion durch einen einzelnen Parameter bestimmt. Zu den meistverwendeten Kerndichten-Schätzern zählen Epanechnikov, Tri-Cubisch und Gaussian. Bei homogenen Beobachtungspopulationen kann zur Beschreibung des Survivalverlaufs die empirische Survivalfunktion direkt angewendet werden. Als nichtparametrisches Verfahren, insbesondere bei Rechtszensierung, eignet sich der Kaplan-Meier-Schätzer.

Zunächst erfolgt die Definition der Risikomenge für jeden Zeitpunkt t aus allen Personen, die zum Zeitpunkt t unter Beobachtung $R(t) = \sum_i 1\{T_i \geq t\}$ standen.

Für den Fall von Rechtszensierungen wird zur Vereinfachung angenommen, dass alle zensierten Ereignisse entweder am Ende des Intervalls stattfinden oder sich zu einem Zeitpunkt ereignen, zu dem nie Sterbefälle stattfinden. Folglich definiert sich das Sterberisiko innerhalb eines betrachteten Zeitintervalls j als bedingte Sterbewahrscheinlichkeit $\hat{p}_j = (1 - d_j / n_j)$ mit allen bis dato noch zu Beginn des Intervalls j noch lebenden Personen n_j und den Sterbefällen d_j , die innerhalb dieses betrachteten Intervalls j stattfinden.

Die Schätzung der Survivalfunktion $\hat{S}(t)$ durch den Kaplan-Meier-Schätzer (vgl. Cleves et al. 2008, S. 93–97, und Klein und Moeschberger 2005, S. 92–93) erfolgt durch

$$\hat{S}(t) = \prod_{j|t_j \leq t} \left(\frac{n_j - d_j}{n_j} \right) \quad (2.5)$$

Diese Funktion verändert sich immer dann, wenn ein Sterbeereignis bzw. eine zensierte Beobachtung an einem bestimmten Zeitpunkt t zu erkennen ist. Um konsistente Schätzergebnisse liefern zu können, muss die Survivalfunktion $S(t)$ gegen null sinken.

Dieses ist immer dann der Fall, wenn der zuletzt beobachtbare Fall nicht zensiert ist. Ist dieses nicht gegeben, dann bleibt $\hat{S}(t)$ für $t > t(n)$ unspezifizierbar, weil für diesen letzten Fall kein exakter Sterbezeitpunkt angegeben werden kann.

Alternativ kann für die Schätzung der kumulierten Hazardfunktion $\Lambda(t)$ der Nelson-Aalen-Schätzer (vgl. Cleves et al. 2008, S. 107–108, sowie Klein und Moeschberger 2005, S. 94) herangezogen werden, der eine vollständige Definition der Sterberisiken auch der zuletzt beobachtbaren Ereignisse sicherstellt.

$$\Lambda(t) = \begin{cases} 0, & \text{wenn } t \leq t_1 \\ \sum_{j|t_j \leq t} \frac{d_j}{n_j}, & \text{wenn } t_1 \leq t \end{cases} \quad (2.6)$$

$$\Lambda(t) = -\ln\{S(t)\}$$

Im Gegensatz zum Kaplan-Meier-Schätzer, der die Schätzung der Survivalfunktion vornimmt, erfolgt mit dem Nelson-Aalen-Schätzer alternativ dafür die Berechnung der kumulativen Hazardfunktion. Der Nelson-Aalen-Schätzer erweist sich als konsistenter, insbesondere bei sehr kleinen Fallzahlen (Andersen et al., 1996).

Des Weiteren sind diese beiden nichtparametrischen Schätzverfahren in gleicher Weise aussagekräftig für die Spezifizierung der Überlebenszeitmodelle und der geschätzten Hazardraten; sie können daher mit relativ geringem technischen Aufwand ausgeführt werden (Klein und Moschberger, 2005).

Bei den nichtparametrischen Verfahren gibt es eine Reihe von Testprozeduren wie log rank, Wilcoxon, Tarone-Ware- oder auch Peto-Peto-Prentice (Klein und Moeschberger, 2005), mit denen Survivalverläufe von zwei oder mehr Vergleichsgruppen auf Gleichheit überprüft werden können. Dabei beziehen sich die Teststatistiken nicht nur auf einen konkreten Zeitpunkt, sondern tätigen einen so genannten „overall“-Vergleich für den gesamten Survivalverlauf.

Dafür werden für jeden Ausfallzeitpunkt die Verhältnisse zwischen den erwarteten und den beobachteten Ereignissen gebildet und mit allen beobachtbaren Ausfallzeiten verbunden. Die einzelnen Testverfahren unterscheiden sich nur bezüglich der Gewichtung der Verhältnisse für die jeweiligen Ausfallzeiten.

2.3. Proportional-Hazards-Modell nach Cox

Proportional-Hazards-Modelle versuchen die Schätzergebnisse für die Ereignisdauer mit beobachteten Kovariablen zu erklären. Ihre Methodik unterscheidet sich nicht sehr von den Prozeduren der gewöhnlichen Regressionsanalyse.

Zugrunde gelegt wird dabei der individuelle Hazard $\lambda(t)$ zum Zeitpunkt oder Alter t und dem Vektor der Kovariablen $X = (X_1, X_2, \dots, X_k)$.

Das Proportional-Hazards-Modell nach Cox (1972) (vgl. Cleves et al. 2008, S. 129 ff., Hosmer et al. 2008, S.69 ff., sowie Wienke 2006, S. 19–20) kann dann spezifiziert werden mit:

$$\lambda(t, X) = \lambda_0(t)g(X) \quad (2.7)$$

wobei $g(X) = g(\beta, X)$ gilt; dann folgt $g(X) = e^{\beta^T X}$

mit $\lambda_0(t)$ als Baseline-Hazard, $X = (X_1, X_2, \dots, X_k)$, dem Vektor für die Kovariablen und $\beta = (\beta_1, \beta_2, \dots, \beta_k)$, dem Vektor für die Regressionskoeffizienten.

Die Baselinefunktion $\lambda_0(t)$ ist gemäß der Modellannahme für alle Individuen der Untersuchungspopulation gleich. Die Kovariablen wirken dabei multiplikativ auf den Baseline-Hazard. Das Charakteristische für dieses Modell ist die Abgrenzung von Alters- bzw. Timing-Effekten, welche die Baselinefunktion betreffen, von den separaten Einflüssen der Kovariablen (Wienke, 2006). Somit wird sichergestellt, dass sich die Hazards sowohl für die Individuen als auch für die gesamte Bevölkerungsgruppe über die Zeit proportional konstant verhalten.

Die Baselinefunktion bleibt von einer Parametrisierung unberührt, deshalb wird das Proportional-Hazards-Modell von Cox auch den semiparametrischen Verfahren zugeordnet.

Die Survivalfunktion bei stetigen Lebenszeiten (Vgl. Wienke, 2006 S.20) kann für das Cox-Modell wie folgt formuliert werden:

$$S(t) = S_0(t) e^{-\sum_{i=1}^k \beta_i X_i}, \text{ und } \beta \text{'s als unbekannte Regressionskoeffizienten}$$

$S_0(t)$ ist eine beliebige Basis-Survivalfunktion und definiert den nichtparametrischen Modellteil, $\exp(\beta_i X_i)$ definiert den Effekt der Kovariablen auf die Hazardfunktion und bildet somit den parametrischen Modellabschnitt. Zwar können für die Baselinefunktion auch parametrische Verteilungen wie z.B. Weibull oder Gompertz-Makeham herangezogen werden, eine genaue Spezifikation ist aber für das Cox-Modell nicht unbedingt notwendig.

Wie schon erwähnt, ist die wichtigste Annahme der Proportional-Hazards-Modelle ein von der Zeit unabhängiges konstantes Verhältnis zwischen den Hazards von zwei unterschiedlichen Beobachtungseinheiten, welche dann nur durch die Regressionskoeffizienten der Kovariablen bestimmt werden. Nur wenn diese Annahme erfüllt ist, kann das Cox-Modell auch sinnvolle Schätzungen für die Hazards bieten.

Dabei kommt die Methode des partiellen Likelihoods zum Tragen, die aber nicht immer eine sinnvolle Alternative des Maximum-Likelihood-Ansatzes der vollparametrischen Regressionsmodelle ist.

Der klare Vorteil dieses Modells liegt zweifelfrei darin, dass für die Baselinefunktion keine spezifische statistische Verteilungsfunktion definiert werden muss.

Dann besteht aber die Gefahr, dass die geschätzte Modellanpassung allein auf die Verteilungsfunktion der Kovariablen zurückgeführt wird, obwohl sich die individuelle Ausfallneigung zwischen den Personengruppen stark unterscheidet.

Dennoch findet das Cox-Hazards-Modell in der wissenschaftlichen Praxis vielseitige Anwendung, obwohl gerade die Regressionskoeffizienten, bedingt durch die Partial-Likelihood-Annahme, oftmals nur schwer interpretiert werden können.

Gemäß der Aussage von Cox (1972, S. 201): „Of course, another issue is the physical or substantive basis for the proportional hazards model. I think that’s one of its weaknesses that accelerated life models are in many ways more appealing because of their quite direct physical interpretation. Particularly in an engineering context“ scheinen die Accelerated-Failure-Time-Modelle (AFT-Modelle) eine bessere Alternative zu sein.

Im Cox-Modell sollen die unbekannten Parameter $\lambda_0(t)$ und β geschätzt werden. Da es sich bei der Baselinefunktion $\lambda_0(t)$ in diesem Fall um einen unendlichdimensionalen Parameter handelt, lässt sich die Maximum-Likelihood-Methode nicht umsetzen. Daher müssen die Regressionsparameter β entsprechend von der Baselinefunktion gelöst werden, damit sie dann unabhängig geschätzt werden können.

Charakteristisch für den partiellen Likelihood, ist die Annahme einer bestehenden Unabhängigkeit von der Baselinefunktion $\lambda_0(t)$, die Funktion wird nur durch die Reihenfolge der Ereignisse bestimmt und zensierte Fälle werden nur über die Risikomenge berücksichtigt.

Partielle Likelihood

Die Kovariablen X_i lassen sich ohne Spezifizierung der Baselinefunktion mit dem partiellen Likelihood (vgl. Cleves et al. 2008, S. 146) schätzen.

$$L(\beta_x) = \prod_{j=1}^k \frac{\exp(X_j \beta_x)}{\sum_{i \in R_j} \exp(X_i \beta_x)}$$

mit R_j als Risikomenge zum Zeitpunkt t_j (2.8)

2.4 Parametrische Verfahren

Im Gegensatz zu den bereits behandelten nichtparametrischen und semiparametrischen Verfahren wird bei den parametrischen Verfahren für die Survivalfunktion $S(t)$, die Hazardrate $\lambda(t)$ und die kumulierte Hazardfunktion $\Lambda(t)$ jeweils eine analytische Verteilungsfunktion formuliert.

2.4.1 Exponentialverteilung

Als eines der einfachsten parametrischen Modelle ist die Exponentialverteilung ($T \sim \text{Exp}(\lambda)$) durch ein über die Zeit konstantes Sterberisiko gekennzeichnet.

Die Wahrscheinlichkeit, innerhalb eines Zeitintervalls zu versterben, wird somit nur über die Länge des betrachteten Intervalls bestimmt. Das bedeutet, dass die Verteilung $T - t$ konditional von $T \geq t$ immer jener Ausgangsverteilung entspricht. Somit wird die Exponentialverteilung als einzige der statistischen Verteilungsfunktionen nicht durch den Zeitpunkt des Ereigniseintritts bestimmt. Alle Parameter λ sind stets positiv, da $\lambda = 1$ gilt.

Parametrische Schätzungen mit der Exponentialverteilung werden im Grunde nur durch einen Parameter bestimmt und reagieren deshalb sehr sensibel, sind aber, bedingt durch diese Einschränkungen, gegenüber anderen statistischen Verteilungsfunktionen nicht sehr flexibel. (vgl. Wienke 2006, S. 7-8., sowie Duchateau und Jansen 2008, S. 29).

$$\text{Dichtefunktion: } f(t) = \lambda e^{-\lambda t}$$

$$\text{Survivalfunktion: } S(t) = e^{-\lambda t}$$

$$\text{Hazardfunktion: } \lambda(t) = \lambda, \quad \lambda > 0$$

$$\text{Kumulative Hazardfunktion: } \Lambda(t) = \lambda t$$

2.4.2 Weibullverteilung

Das Weibull-Modell gilt immer dann als sinnvolle Erweiterung der Exponentialverteilung, wenn eine Schätzung mit zwei Parametern erfolgt. Insbesondere durch die Einführungen des zweiten positiven Parameters erhält das Modell seine maximale Flexibilität.

Gerade wegen seiner doch sehr einfachen Struktur und doch relativ leichten mathematischen Umsetzung ist dieses Modell bei verschiedenen Fragestellungen aus unterschiedlichen Fachbereichen sehr beliebt (vgl. Wienke 2006, S. 8, sowie Duchateau und Jansen 2008, S. 29).

$$\text{Dichtefunktion: } f(t) = \lambda \gamma t^{\gamma-1} e^{-\lambda t^\gamma}$$

$$\text{Survivalfunktion: } S(t) = e^{-\lambda t^\gamma}$$

$$\text{Hazardfunktion: } \lambda(t) = \lambda \gamma t^{\gamma-1}$$

$$\text{Kumulative Hazardfunktion: } \Lambda(t) = \lambda t^\gamma$$

Für den Fall $\gamma = 1$ mit einer konstanten Ausfallrate ist die Exponentialverteilung gegenüber der Weibull-Verteilung vorzuziehen. Sobald aber mit einer steigenden ($\gamma > 1$) oder fallenden ($\gamma < 1$) Ausfallrate zu erwarten ist, muss das Weibull-Modell beachtet werden.

Die Weibull-Hazards konnten theoretisch von Pike (1966) für verschiedenste Krebserkrankungen nachgewiesen werden; ob sie sich auch auf andere Erkrankungen anwenden lassen, ist bisher noch nicht spezifiziert worden.

Sicher ist aber, dass ein Weibull-Modell nur für jene Hazards sinnvolle Schätzungen liefert, die nicht unimodal oder badewannenförmig sind (Wienke, 2006).

2.4.3 Gompertz – Verteilung

Dieses Modell kennzeichnet sich durch die zwei Parameter (a, b) mit einer über die Zeit monoton-steigenden Ausfallrate. Wenn $(b = 0)$, dann ergibt sich daraus als Spezialfall die Exponentialverteilung. Die Gompertz-Verteilung ist insbesondere bei demographischen und biologischen Fragestellungen sehr verbreitet (vgl. Wienke 2006, S. 10, sowie Duchateau und Jansen 2008, S. 29).

$$\text{Dichtefunktion: } f(t) = ae^{bt} e^{-\frac{a}{b}(e^{bt}-1)}$$

$$\text{Survivalfunktion: } S(t) = e^{-\frac{a}{b}(e^{bt}-1)}$$

$$\text{Hazardfunktion: } \lambda(t) = ae^{bt}$$

$$\text{Kumulative Hazardfunktion: } \Lambda(t) = \frac{a}{b}(e^{bt} - 1)$$

Die Hazardfunktion wächst von a zum Zeitpunkt t_0 bis ∞ zum Zeitpunkt ∞ .

Durch die Erweiterung mit einer Konstante c lässt sich der Gompertz-Makeham-Ansatz bilden.

$$\lambda(t) = ae^{bt} + c$$

Der Gompertz-Makeham-Ansatz setzt sich demnach aus einer Summe aus einer altersunabhängigen Konstanten c und einer altersabhängigen Komponente ae^{bt} zusammen, die mit dem Alter exponentiell ansteigt. Daher kann insbesondere sehr gut die Sterblichkeit im mittleren bis hohen Lebensalter (zwischen 30 und 80 Jahren) beschrieben werden.

2.4.4 Gamma-Verteilung

Die Gamma-Verteilung ist ein Sonderfall der Exponentialverteilung (vgl. Wienke 2006, S. 11). In der Survivalanalyse kann aber dieses Modell nur bedingt eingesetzt werden, weil die Survival- und die Hazardfunktion ein unvollständiges Gamma-Integral enthalten und somit niemals eine festdefinierte Obergrenze erreichen können.

$$I_k(t) = \frac{\int_0^t s^{k-1} e^{-s} ds}{\Gamma(k)} \quad (2.9)$$

Bei den Maximum-Likelihood-Schätzungen können dann Verzerrungen bei den Parameterschätzungen eintreten, da die dafür benötigte vollständige Bestimmung des Gamma-Integrals nicht möglich ist.

Die zufällige Variable T ist gammaverteilt mit dem Form-Parameter k und dem Skalen-Parameter λ ($T \sim \Gamma(k, \lambda)$)

$$\text{Dichtefunktion: } f(t) = \frac{\lambda^k t^{k-1} e^{-\lambda t}}{\Gamma(k)} \quad k, \lambda > 0$$

$$\text{Survivalfunktion: } S(t) = 1 - I_k(\lambda t)$$

$$\text{Hazardfunktion: } \lambda(t) = \frac{\lambda^k t^{k-1} e^{-\lambda t}}{(1 - I_k(\lambda t))\Gamma(k)}$$

$$\text{Laplace-Transformation: } L(s) = Ee^{-Ts} = \left(1 + \frac{s}{\lambda}\right)^{-k}$$

Sobald $k = 1$ gilt, wird die Gamma-Funktion in eine Exponentialverteilung umgewandelt.

3. Der Einfluss von unbeobachteter Heterogenität in der Mortalitätsforschung

Bei der Analyse von Überlebenszeiten in Populationen werden bei vielen epidemiologischen Studien nur wenige Kovariablen wie Alter und Geschlecht berücksichtigt.

Andere Regressoren wie das individuelle Gesundheitsverhalten, die Lebensführung, das Rauchverhalten oder auch genetische Einflüsse können ebenfalls zu wichtigen Erklärungsgaranten werden, wobei in vielen Untersuchungen oftmals der Informationsgehalt dieser Faktoren auf der Individualebene zu gering ist und somit deren Relevanz oftmals unterschätzt wird (Westerman et al., 2009).

Auch Individuen können sich bezüglich ihrer Eigenschaften stark unterscheiden. Unterschiede bezüglich der beobachtbaren Merkmale wie beispielsweise Alter und Geschlecht können oftmals mit geringem Aufwand direkt ermittelt werden. Bei der Betrachtung nicht direkt beobachtbarer Merkmale wie individueller Einstellungen zur Lebensführung oder nicht physisch messbarer Gesundheitsmerkmale erweist sich eine direkte Identifikation als sehr schwierig, zumal sich diese Faktoren im Laufe eines Lebens auch noch verändern können. Es besteht eine Vielzahl von Diskrepanzen zwischen den individuellen Empfindungen und den tatsächlich beobachtbaren Zuständen und der Problematik der objektiven Messung und der numerischen Bewertung von Einstellungen.

Beispielhaft sind hierbei die ermittelten geschlechtsspezifischen Unterschiede in der Lebensqualität von Typ-2-Diabetikern im Disease-Management-Programm (DMP).

Mit dieser Studie konnten scheinbar subjektive Unterschiede (Miksch et al., 2008) in der Lebensqualität von Männern und Frauen beobachtet werden. Für Männer konnten gegenüber den Frauen in allen Skalen des SF-36 höhere Werte ermittelt werden. Bei Differenzierung zwischen DMP- und Nicht-DMP-Teilnehmern hatten die Männer bei der Nicht-DMP-Teilnahme höherer Werte in der Zufriedenheit, dagegen waren bei den Frauen nur die Werte der DMP-Teilnehmerinnen erhöht.

Die Autoren begründen diese vermuteten geschlechtsspezifischen Unterschiede in der Bewertung der Lebensqualität damit, dass Personen, die an Disease-Management-Programmen teilnehmen, sich intensiver mit ihrer Krankheit auseinandersetzen und somit eher bereit sind ihre bisherigen Lebensgewohnheiten zu ändern. Es zeigt aber auch, dass Männer und Frauen auf die Krankheit unterschiedlich reagieren.

Männliche Patienten empfinden die Symptome bzw. die Krankheitsursachen als rein körperlich, Frauen bewerten ihre Krankheitsbilder häufiger psychosozial. Eine intensivere Auseinandersetzung mit der Krankheit könnte demnach bei Männern dazu führen, dass in Abhängigkeit von persönlichen Bewertungen sich die Lebensqualität verschlechtern könnte.

Bei Frauen könnte dagegen eine beobachtbare Mehrbeschäftigung mit der Krankheit bewirken, dass die individuellen Empfindungen einer verbesserten Versorgung eher zunehmen und daher mit einer höheren Lebensqualität einhergehen.

Zwischen den beobachtbaren und nichtbeobachtbaren Risikofaktoren besteht oft eine enge Beziehung, die sehr häufig durch eine gegenseitige Bedingtheit gekennzeichnet ist. Beispielsweise könnte sich mit zunehmendem Lebensalter das individuelle Gesundheitsbewusstsein positiv oder negativ verändern. Ebenso könnte auch ein allgemein als gut bewerteter Gesundheitszustand sich positiv bzw. negativ auf das Lebensalter auswirken.

Daher müssen für die sukzessive Datenanalyse verschiedene Arten von Variabilität für die behandelten Risikofaktoren (Wienke, 2006) berücksichtigt werden: 1. die theoretisch messbare Variabilität für direkt beobachtbare Risikofaktoren und 2. die Varianzen von jenen Risikofaktoren, die nicht direkt ermittelt werden können, aber auf andere Risikofaktoren indirekt einwirken können, sodass sich der Effekt von unbeobachteter Heterogenität aufzeigen lässt.

Eine Berücksichtigung von unbeobachteter Heterogenität ist immer dann sinnvoll, wenn die empirischen Schätzverfahren zu unerwarteten oder sogar zu einander widersprechenden Ergebnissen führen (Vaupel et al., 1979; Yashin et al., 2008).

Bestimmte Teilpopulationen von Individuen können einem höheren Risiko häufiger ausgesetzt sein und können durch die feststellbare Selektion vorzeitig ausscheiden, was dann in einer selektierten Subgruppe von Individuen mit geringen Risiken resultiert (Keyfitz und Caswell, 2005).

Mit anderen Worten: Der Effekt von unbeobachteter Heterogenität kann dazu führen, dass sich Teilpopulationen von Hochrisikofällen im Laufe der Beobachtungszeit zu Individuen mit geringeren Lebensrisiken entwickeln können und umgekehrt. Diese Feststellung zeigt erneut auf, dass Populationen sich nicht immer aus rein homogenen Personengruppen mit identischen Lebensrisiken zusammensetzen.

3.1 Die Wirkung von unbeobachteter Heterogenität an einem rechnerischen Beispiel

An dem nun folgenden fiktiven Beispiel soll die Wirkung von unbeobachteter Heterogenität innerhalb von Teilpopulationen aufgezeigt werden. Es soll die Sterblichkeit von zwei unterschiedlichen Personengruppen in Abhängigkeit vom Lebensalter verglichen werden.

Die Mitglieder der einen Teilpopulation (Behandlungsgruppe) erfahren eine medizinische Intervention, welche maßgeblich eine Reduktion der Gesamtsterblichkeit herbeiführen soll. Die Behandlung der Probanden in der Vergleichsgruppe erfolgt ohne diese medizinische Intervention.

Der Sterblichkeitsvergleich beschränkt sich aus Vereinfachung auf 6 Altersgruppen, eine Untersuchung realer Populationen würde eine weitaus größere Anzahl von Altersstufen einschließen. Beide Teilpopulationen sollen sich bezüglich ihres Durchschnittsalters und in ihrer Sterblichkeit unterscheiden.

Die Subgruppe mit der medizinischen Intervention hat ein Durchschnittsalter von ca. 62,9 Jahren, die Teilpopulation ohne medizinische Intervention ist mit 58,6 Jahren um mehr als vier Jahre jünger. Die altersspezifische Sterblichkeit soll für 30 Zeiteinheiten konstant sein, wobei die Werte für die Treatment-Gruppe insgesamt doppelt so hoch sind wie für die Vergleichsgruppe. Die Sterblichkeit in der Teilpopulation der Behandlungsgruppe soll sich stärker auf die oberen Alterstufen konzentrieren.

Tabelle 4: Die Wirkung von Selektion in unterschiedlich zusammengesetzten Teilpopulationen anhand eines fiktiven Beispiels

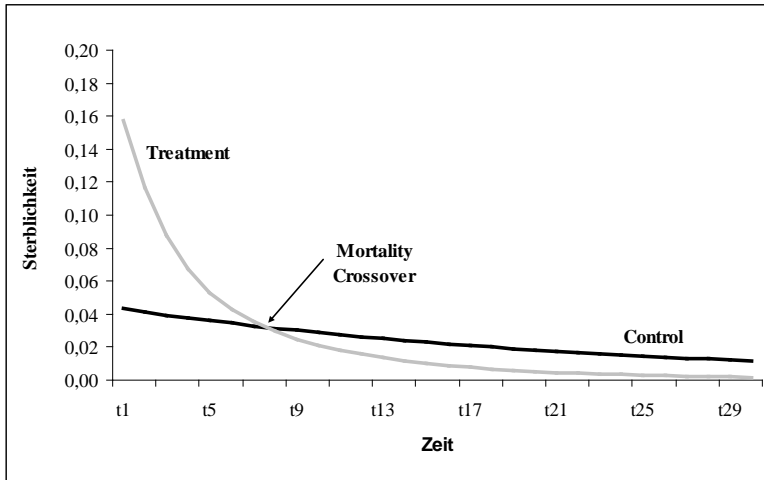
Alter Treatment	absolute Häufigkeit	alterspezifische Sterblichkeit	Alter Control	absolute Häufigkeit	altersspezifische Sterblichkeit	Durchschnitts- alter Treatment	Durchschnitts- alter Control
43	90	0,0001	47	108	0,05	63,0	58,6
45	95	0,0001	49	89	0,05		
53	44	0,0005	59	68	0,001		
68	71	0,1	63	138	0,001		
73	112	0,21	70	102	0,1		
87	103	0,49	88	10	0,2		
Gesamt	515		Gesamt	515			

Eigene Berechnung

Es werden nun die Sterblichkeitsverhältnisse für beide Vergleichsgruppen im Zeitablauf verglichen. Wie deutlich zu erkennen (siehe Abb. 1) ist, werden in der Kontrollgruppe mit der deutlich geringeren Sterblichkeit die so genannten schlechten Risiken wesentlich schwächer ausselektiert, was auf die geringere altersspezifische Sterblichkeit in den oberen Altersgruppen zurückzuführen ist. In der Treatmentgruppe erfolgt die Selektion wesentlich stärker, weil zum einen die Mortalität insgesamt doppelt so hoch ist, zum anderen die Altersstruktur wesentlich älter ist.

Diese innere Selektion schwächt sich im Laufe der Beobachtungszeit ab, da die Sterblichkeitsreduktion auch in der Kontrollgruppe einsetzt und dadurch das Durchschnittsalter ansteigt. Ebenso ist zu erkennen, dass im Gruppenvergleich die Selektion innerhalb einer jeden Teilpopulation nach unterschiedlichen Gesetzmäßigkeiten stattfindet. Es kann dazu führen, dass sich die Sterblichkeiten beider Vergleichsgruppen überschneiden und dann in einem Mortality Crossover (Coale und Kisker, 1986; Dupre et al., 2006; Elwert und Wrigley-Field, 2009; Lynch et al., 2003) resultieren, was sehr häufig ein Anzeichen von unbeobachteter Heterogenität ist. Je länger die Teilpopulationen unter Beobachtung stehen, desto eher scheinen sich die Sterblichkeitsunterschiede anzugleichen, was unmittelbar auf die Abnahme des Selektionsdruckes in beiden Selektionsgruppen zurückzuführen ist.

Abbildung 2: Darstellung der kumulierten Sterblichkeit für
die Interventions- und die Kontrollgruppe



Eigene Darstellung

Beide Teilbevölkerungen enthalten nun einen wesentlich höheren Anteil von Individuen mit einer deutlich geringeren Sterblichkeitsneigung, da jene Fälle mit der höheren Mortalitätseigenschaft schon relativ frühzeitig als Hochrisiko-Fälle ausselektiert wurden.

Folglich wird sich die Zusammensetzung der Teilbevölkerungen, dabei insbesondere der Treatmentgruppe, von einer ehemals ausgeprägten Mortalität von Hochrisikofällen im oberen Lebensalter zu einer weniger intensiven Sterblichkeit von Individuen im mittleren Lebensalter entwickeln.

Biologische Selektion in unterschiedlichen Beobachtungsgruppen ist immer dann feststellbar, sobald Individuen sich bezüglich ihrer lebensdauerspezifischen Eigenschaften unterscheiden. Diese feststellbaren Selektionsmechanismen folgen für die einzelnen Geburtsjahrgänge scheinbar festgelegten Mustern. Für zwei zu vergleichende Subgruppen mit jeweils unterschiedlich starkem Selektionsdruck wird für das Selektionsmerkmal eine empirische Normalverteilung (interne Homogenität) unterstellt (Manton et al., 1979). Diese Annahme ist deshalb notwendig, um sicherzustellen, dass andere Faktoren, die einen Einfluss auf die genetische Varianz in der Population haben, die Selektion nicht verzerren.

3.2 Identifikation eines Mortality-Crossovers

Aber die Identifikation eines Mortality-Crossovers ist auch dann möglich, wenn bei nur einer Teilpopulation das jeweilige zu untersuchende Merkmal normalverteilt ist und sich diese Gruppe durch interne Homogenität kennzeichnet. In der anderen Vergleichsgruppe ist aber eine interne Heterogenität zu erkennen, die durch eine hohe Varianz innerhalb des Selektionsmerkmals wie z.B. eine vorzeitig erhöhte Mortalität und eine aufgeschobene reduzierte Sterblichkeit bestimmt wird. Die Exaktheit eines Selektions-Modells für die fundierte Erklärung eines Crossovers hängt nicht nur allein von der Merkmalsverteilung innerhalb der betrachteten Bevölkerung ab.

Sobald der Crossover allein nur durch die Selektion innerhalb des Modells erklärt wird, muss die Teilpopulation mit der zu Beginn beobachteten erhöhten Sterblichkeit im Hinblick auf eine höheren Lebensdauer schneller homogener werden als die Teilbevölkerung mit geringem Selektionsdruck. Zu erwähnen sind dabei zwei wichtige Situationen (Manton et al., 1979), bei denen trotz einer unterschiedlich intensiven Selektion kein Crossover festzustellen ist:

1. Wenn die festgestellten Selektionsunterschiede zwischen den beiden Untersuchungsgruppen zu gering sind, können durchaus beide Teilpopulationen sich relativ schnell zu homogenen Einheiten entwickeln, ohne dass dabei ein Crossover zu erkennen ist.

Unter der Annahme, dass genetische Faktoren die Heilungschancen einer bestimmten Krankheit X verbessern können, haben Individuen, welche diese besonderen genetischen Eigenschaften besitzen, nicht nur bessere Chancen, schneller zu gesunden, sondern gleichzeitig verlängert sich dadurch möglicherweise auch ihre Lebensdauer.

Der dadurch ermittelte Vorteil bezüglich der Lebensdauer kann daher keinen sicheren Erklärungsgehalt für den vermuteten Crossover von zwei genetisch sich unterscheidenden Gruppen darstellen, weil nicht unterschieden werden kann, ob sich der genetische Vorteil direkt nur auf die Krankheit bezieht oder auf die Lebenserwartung auswirkt.

2. Wenn der Selektionsdruck zu hoch ist und relativ gleichmäßig über alle Altersgruppen der „benachteiligten Gruppe“ verteilt ist, Folgt daraus, dass nur eine Selektion in Teilpopulationen immer nur dann ein Crossover erzeugt, wenn sich die Mortalität und die Varianz der Selektionsmerkmale für die Lebensdauer zwischen zwei Teilgruppen in einer scheinbar festgelegten Gesetzmäßigkeit verlaufen.

Ebenso können individuelle Unterschiede bezüglich der genetischen Eigenschaften oder der Lebensdauer durch verschiedene exogene Umweltfaktoren wie beispielsweise die Art der Lebensführung, das Rauchverhalten, die Ernährung oder berufliche Belastungen hervorgerufen werden.

Die lebensdauerspezifischen Eigenschaften einer Bevölkerung werden demzufolge von der nicht zufälligen Sterblichkeit (Selektion) und den exogenen Faktoren des Alterns beeinflusst. In Populationen, die durch intensive Sterblichkeit und homogene Umweltfaktoren gekennzeichnet sind, wird die Selektion allein durch die Sterblichkeit bestimmt.

Sollte aber die Sterblichkeit nur sehr gering ausgeprägt sein, aber die exogenen Faktoren sich heterogen verteilen, dann erfolgt die Selektion allein durch Varianz der äußeren Umweltfaktoren.

Zusammenfassend lässt sich Folgendes festhalten: Die Existenz eines Crossovers von Sterblichkeiten wird bestimmt durch zwei wesentlichen Faktoren.

1. Es müssen sich die Populationen bezüglich ihrer Sterblichkeiten deutlich unterscheiden, d.h., ein Sterblichkeitsunterschied muss feststellbar sein.
2. Es müssen sich die Populationen bezüglich ihrer intragruppenspezifischen Zusammensetzung auch deutlich unterscheiden, d.h., in einer der beiden Teilgruppen sollten wesentlich mehr ungünstige Fälle vorhanden sein, welche dann einen höheren Selektionsdruck auslösen. Die Effekte und Einflüsse eines solchen Mortalitäts-Crossovers lassen sich anhand der Sterblichkeitsunterscheide zwischen der weißen und „african american“ US-Bevölkerung empirisch belegen (Elo und Preston, 1997; Kitigawa und Hauser, 1973; Preston et al., 1996; Roger, 1992).

4. Die statistische Überprüfung von unbeobachteter Heterogenität

Die Standardverfahren der Survivalanalyse wie der nichtparametrische Kaplan-Meier-Schätzer sowie das Proportional-Hazards-Modell unterstellen stets eine homogene Untersuchungsbevölkerung. Alle Individuen haben ein identisches Risiko, gemäß dem Hazard $\lambda(t)$ aus der Population auszuschneiden, oder sie besitzen die gleiche Sterbefunktion $S(t)$. Einige Modelle wie z.B. das klassische Cox-Modell können nur mit Hilfe von bekannten Kovariablen-Eigenschaften sinnvolle Schätzergebnisse liefern, ohne dabei sicherzustellen, ob diese Modelle alle Erklärungsfaktoren enthalten oder mögliche Effekte nur scheinbar erklären. Die nun folgenden methodischen Ansätze können direkt die Auswirkung von unbeobachteter Heterogenität erstellen.

4.1 Das Problem der unbeobachteten Heterogenität

Viele epidemiologische Studien berücksichtigen bei der Analyse der Sterblichkeitsverhältnisse von Populationen nur wenige Kovariablen wie beispielsweise Alter und Geschlecht. Daneben können weitere wichtige Einflusskomponenten wie der Gesundheitszustand, die allgemeine Lebensführung, insbesondere das Rauchverhalten oder die Arbeitsbelastungen, aber auch genetische Faktoren für die feststellbaren Unterschiede in der Sterblichkeit von verschiedenen Subgruppen mit verantwortlich sein.

Dabei ist es oftmals sehr schwierig, bei der Modellbildung alle wichtige Risikofaktoren zu berücksichtigen, zumal oftmals die Informationen auf der Individualebene dieser Parameter zu gering oder gänzlich nicht vorhanden sind, sodass die Relevanz dieser Faktoren dann häufig unterschätzt werden kann.

Insbesondere klinische Studienbevölkerungen weisen häufig keine homogenen Strukturen auf. Das zeigt sich z.B. dadurch, dass bestimmte Therapiemaßnahmen bei einigen Patienten effizienter wirken als bei anderen Probanden. Ebenso können bei der medikamentösen Therapie auch unterschiedliche körperliche Voraussetzungen wie das Geschlecht, die Körpergröße oder auch das Gewicht mit entscheidend bei der dosis-abhängigen Verordnungen sein und damit einhergehend sich dann entweder positiv oder negativ auf den Krankheitsverlauf auswirken.

Auch die Heilungschancen für bestimmte Krankheiten können sich im Laufe eines Lebens ändern. Insbesondere bei vielen Krebserkrankungen, so z.B. bei Ovarienkarzinomen und ihrer Diagnose, ist die Bestimmung des Tumorstadiums für den weiteren Prognoseverlauf entscheidend.

In dem frühen Stadium IA (ein Ovarium ist befallen, die Kapsel ist aber intakt) der FIGO-Stadieneinteilung liegen die 5-jährigen Überlebensraten noch bei 90%, im Stadium IIA (Ausdehnung des Tumors auf den Uterus und den Tuben) sinken diese auf 60%, im Stadium III (Anwachsen des Tumors auf mehr als 2cm außerhalb des Beckens und Bildung von Metastasen) beträgt sie dann 30% und für das letzte Stadium nur noch 10% (Medical Tribune, 2009; Trimbo et. al., 2010).

Sehr häufig wird ein gleichzeitiger Befall beider Ovarien diagnostiziert, verbunden mit einem raschen Wachstum der Tumoren, wobei sich die Wachstumsgeschwindigkeiten innerhalb der Patientinnen sehr stark unterscheiden können.

Wird eine so genannte First-Line-Chemotherapie (nach Schema Carboplatin/Pacitaxel mit Taxol 175mg/m² KG, Carboplatin AUC2) wie gewöhnlich für eher fortgeschrittene Ovarialkarzinom-Stadien auch schon in den frühen Stadien I–II angewendet, weil aufgrund der Komplexität der Krankheitssymptomatik die Art des Stadiums nicht eindeutig identifiziert konnte, ist es durchaus möglich, dass diese Therapieform für diese Patientengruppe überhaupt nicht effizient und daher eine Änderung des Therapieplans erforderlich ist.

4.2 Numerische Tests zur Überprüfung der Proportional-Hazards-Bedingung

In Abschnitt 2.3.1 erfolgte bereits eine ausführliche Darstellung des Proportional-Hazards-Ansatzes von Cox (1972). Die Reliabilität der Schätzergebnisse wird dabei durch zwei wichtige Annahmen bestimmt, die nicht verletzt werden dürfen (NG'Andu, 1997):

1. eine lineare Beziehung zwischen den log hazards bzw. kumulierter log hazard und den erklärenden Kovariablen;
2. zeitliche Unabhängigkeit der Kovariablen in Bezug auf die Hazards-Funktion; das bedeutet, dass sich die Hazards-Ratios von zwei verschiedenen Individuen nicht über die Zeit verändern dürfen.

Für die Überprüfung der Gültigkeit der PH-Bedingung haben sich zahlreiche numerische Tests, aber auch graphische Methoden etabliert, die je nach wissenschaftlicher Fragestellung, den Modellannahmen sowie der Zensierung oder Nichtzensierung sehr unterschiedliche Anpassungen vornehmen (O'Quigley und Pessione, 1989; Choi und Park, 2007).

Cox (1972) hatte bereits in seiner Originalarbeit mit der Einführung von zeitabhängigen Variablen zur Modellierung von Interaktionseffekten⁶ eine Methode entwickelt, durch die sich die PH-Bedingung sehr effizient überprüfen lässt. Diese Modellerweiterung erlaubt somit die Behandlung von zeitabhängigen Effekten sowie deren Überprüfung auf statistische Signifikanz.

Eine weitere methodisch sehr simple Prozedur zur Überprüfung der PH-Bedingung ist die PHGLM-Methode von Harrell (1986), welche wiederum auf Schoenfelds Partial-Residual-Test aufbaut (Schoenfeld, 1980, 1982). Dabei wird zur Vergleichbarkeit eine Fisher's z-Transformation der Pearson-Korrelation zwischen den partiellen Residuen und den Rank-Koeffizienten der Ereigniszeit vorgenommen (NG'Andu, 1997). Zur Vereinfachung sind die Residuen zeitunabhängig und beeinflussen daher in keiner Weise die Koeffizienten der Baseline-Hazards.

Für die Teststatistik gilt daher: $Z = \rho \sqrt{(n_u - 2) / (1 - \rho^2)}$, mit ρ als Korrelation zwischen den Residuen und der Ereigniszeit und n_u , der Zahl der unzensierten Beobachtungen. Die Werte der Teststatistik können sowohl positiv als auch negativ sein: positiv immer dann, wenn die Hazards-Ratios über die Zeit größer werden, negativ, wenn sie über die Zeit sinken.

⁶ Ein zeitabhängiger Interaktionseffekt kann beispielsweise durch $\text{sex} * \log(t)$ definiert werden.

Ein sehr verbreitetes und etabliertes Verfahren zur Überprüfung der Proportional-Hazard-Bedingung sind Score-Tests für die Schoenfeld-Residuen (Grambsch und Therneau, 1994; Thernau et al., 1990). Für jede einzelne Kovariable eines jeden Individuums i wird separat ein Residuum ausgewiesen.

Die Schoenfeld-Residuen werden gebildet aus der Differenz des Kovariablen-Vektors der Individuen i , welche zum Zeitpunkt t aus der Risikobevölkerung ausgeschieden sind, und den erwarteten mittleren Kovariablenwerten für alle Individuen, die zu diesem Zeitpunkt t dem Risiko ausgesetzt sind (Winnett und Sasieni, 2001).

$$\hat{r}_{(i)} = X_{(i)} - \frac{\sum_{j \in \mathcal{R}(i)} X_j \exp(\hat{\beta}^T X_j)}{\sum_{j \in \mathcal{R}(i)} \exp(\hat{\beta}^T X_j)} \quad (4.1)$$

Allerdings ist die Anwendung der Schoenfeld-Residuen (vgl. Winnett und Sasieni, 2001, S. 566) immer dann problematisch, sobald die Risikopopulation nicht vollständig bestimmt werden kann, da einzelne Fälle zensiert sind.

Aus diesem Grund werden so genannte Scaled-Schoenfeld-Residuen verwendet, weil diese nicht nur die Ausfälle, sondern alle Beobachtungen einbeziehen. Die statistische Überprüfung erfolgt durch die Maximierung der partiellen Likelihood-Schätzungen der Kovariablen.

Die Score-Tests für die Schoenfeld-Residuen bieten aus methodischer Sicht viele praktische Vorteile:

1. Bei Modellerweiterung mit neuen Kovariablen können diese an die bestehenden Modellbedingungen angepasst werden, ohne dabei die Ausgangssituation des Modells zu verändern.
2. Es lässt sich ebenso feststellen, welchen Einfluss jedes einzelne Individuum auf die Modellschätzung hat.
3. Die Genauigkeit des Modells kann anhand der Ausfallrate für einen gegebenen konkreten Fall ausgewiesen werden und somit lässt sich die Beeinflussung durch Verzerrungen von Ausreißern minimieren.
4. Es erfolgt eine Überprüfung der Gültigkeit der Proportional-Hazards-Annahme sodass diagnostiziert werden kann, ob die Unveränderlichkeit des Hazard-Raten-Verhältnisses zeitunabhängig gilt

4.3. Die graphische Überprüfung der Proportional-Hazards-Bedingung

Die Gültigkeit der Proportionalitätsannahme kann auch durch graphische Methoden veranschaulicht und geprüft werden. Graphische Verfahren sind durchaus sehr praktikabel, da sie eine separate Betrachtung einzelner Kovariablen erlauben. So argumentierten auch einige Autoren wie Klein und Moesberger (1997), wobei eine korrekte Modellanpassung unter der Proportional-Hazards-Bedingung nur für einzelne wenige Erklärungsgrößen gegeben ist und somit viele numerische Tests nur in Abhängigkeit von der Stichprobengröße ausreichende statistische Power zeigen.

In einer wissenschaftlichen Arbeit von Hess (1995) wurden insgesamt acht verschiedene graphische Verfahren beschrieben, die doch zum Teil von sehr unterschiedlichen Modellannahmen ausgehen. Wenn für die Überprüfungen der Proportionalitätsbedingungen graphische Prozeduren genutzt werden, sollte dem Nutzer bewusst sein, dass die Interpretation der Ergebnisse sehr stark von der Wahl der Darstellungsmethode abhängt, die eher nach subjektiver Wahrnehmung des statistischen Problems als nach allgemeingültigen Auswahlkriterien erfolgt (Persson et al., 2007).

Zur Veranschaulichung werden nur fünf der gebräuchlichsten Methoden zusammenfassend dargestellt:

1. Die Überlebenskurven und die gruppenspezifischen Koeffizienten des Kaplan-Meier-Schätzers:

Die Überprüfung auf Verletzung der Proportional-Hasards-Annahmen erfolgt durch einen Vergleich der Survivalverläufe des Cox-Modells und des unabhängigen nichtparametrischen Modells (Kaplan-Meier-Produkt-Schätzer) separat für jede Untersuchungsgruppe.

Dabei bestimmt die absolute Differenz zwischen den beiden Survivalkurven die Intensität der Verletzung für die PH-Bedingung.

2. Kumulative Baseline-Hazards für verschiedene Gruppen (Andersen-Plot):

Das Verhältnis der kumulativen Baseline-Hazards für zwei verschiedene Untersuchungsgruppen wird im Zeitablauf nur durch eine gemeinsame Linie beschrieben (Andersen, 1982; Andersen et al., 1993). Dabei gilt: Je eher diese Kurve dem linearen Modell ähnelt, desto wahrscheinlicher ist es, dass die PH-Annahme nicht verletzt wurde.

3. Differenz der logarithmierten kumulativen Baseline-Hazards gegenüber der Zeit:

Gemäß der PH-Bedingung wäre die Differenz der logarithmierten kumulativen Hazards über die Zeit konstant (Schumacher, 1990)

4. Geglättete logarithmierte Baseline-Hazards gegenüber die Zeit:

Ähnliche methodische Vorgehensweise wie im dritten Verfahren, allerdings werden Hazards durch ein Glättungsverfahren angepasst.

5. Skalierte Schoenfeld-Residuen (Schoenfeld, 1980):

Bei Gültigkeit der PH-Bedingung sind partielle Residuen zeitunabhängig. Im Plot werden die horizontalen Bänder für die Residuals angezeigt, die um die Log der kumulierten Baseline-Hazards variieren.

4.4 Der theoretische Hintergrund der Frailty-Modelle

Die unterschiedliche Wirksamkeiten von gleichen Therapieformen bei doch vermeintlichen identischen Patientengruppen beweist, dass die Existenz von Heterogenität oftmals nur sehr schwierig identifiziert werden kann.

Unter diesen Umständen haben sich die so genannten Frailty-Modelle entwickelt, welche die Wirkungen von unbeobachteter Heterogenität überprüfen wollen. Die Idee dieser Modelle scheint zunächst sehr trivial: Individuen unterscheiden sich bezüglich ihrer Anfälligkeitswahrscheinlichkeiten gegenüber bestimmten Risikofaktoren.

Jene, die eine höhere Anfälligkeit aufweisen, werden eher aus der Risikobevölkerung ausscheiden bzw. versterben früher und bewirken daher eine systematische Selektion von robusteren Individuen.

Es ist durchaus vorstellbar, dass sich eine Population mit einem zunächst relativ hohen Anteil von Hochrisikofällen durch den Einfluss von Selektionsmechanismen zu einer Personengruppe mit deutlich geringeren Lebensrisiken entwickelt. Eine Nichtberücksichtigung der Effekte (Hougaard, 1991) der unbeobachteten Heterogenität durch die Frailty-Varianz könnte unerwartete Ergebnisse oder alternative bzw. kontroverse Interpretationen erzeugen, wie z.B., dass scheinbare Risikofaktoren, denen eigentlich schädliche Eigenschaften zugeordnet wurden, sich nun plötzlich positiv auf die Prävalenz einer Krankheit auswirken.

Das könnte zur Folge haben, dass Patienten mit chronischen Herz-Kreislauf-Erkrankungen, die während ihres gesamten Lebens geraucht haben, nun empfohlen wird, nicht mit dem Rauchen aufzuhören, weil sich dieses nun präventiv auf das Krankheitsbild auswirkt.

Die unterschiedlichen Wirkungsgrade der Selektionsfaktoren können durch die Mixed Models Beachtung finden, d.h., die Populationen setzen sich aus einer gemischten Struktur von Individuen zusammen, die unterschiedlich hohe Lebensrisiken aufweisen. Diese nicht direkt beobachtbaren Risiken können als „Mixture“-Variable dem so genannten Frailty dargestellt werden. Diese Frailties sind Zufallsvariablen, die einer bestimmten statistischen Verteilung folgen wie binäre, Gamma- und Log-normal-, welche dann sowohl quantitative als auch qualitative Unterschiede (Wienke, 2006) aufzeigen.

Das klassische Frailty-Modell basiert auf der schon bereits dargestellten Proportional-Hazards-Bedingung und wird konditional durch den Zufallseffekt der Frailty bestimmt (Vaupel et al., 1979). Genauer gesagt, der Hazard eines Individuums wird durch eine zufällige, altersunabhängige Variable (Frailty) Z definiert, welche dann direkt auf die Baselinefunktion $\lambda_0(t)$ einwirkt (vgl. Hosmer et al. 2008, S. 297 ff., sowie Wienke, 2006 S. 22 ff.).

$$\lambda(t, Z) = Z\lambda_0(t) \tag{4.2}$$

Die Frailty Z ist somit eine rein zufällige Mixture-Variable, die innerhalb der Untersuchungspopulation variiert. Es wird unterstellt, dass für alle Individuen ein gemeinsamer Faktor existiert, welcher sich möglicherweise direkt auf den Baseline-Hazard auswirkt.

Die Frailty-Verteilung ist dabei standardisiert mit $EZ = 1$. Entscheidend für den Einfluss der Heterogenität in Populationen ist somit die Frailty-Varianz $\sigma^2 = V(Z)$.

Wenn der Skalenwert für die Varianz sehr gering ist, dann werden die Werte von Z sich wohl eher an 1 annähern. Mit steigenden Koeffizienten für σ^2 werden die Werte von Z deutlich größer als 1 sein und der individuelle Hazard $Z\lambda_0(t)$ wird dann stärker durch den Heterogenitätsgrad beeinflusst.

Eine ausführliche mathematische Beweisführung zur Gültigkeit dieser Modellannahmen erfolgt in der Habilitationsschrift von Wienke, 2006, auf den Seiten 23 ff.

4.5 Statistische Verteilungsfunktionen für den Frailty

4.5.1 Gamma-Frailty

Wie bereits erwähnt, setzt sich eine Population aus einer Vielzahl von Individuen zusammen, die sich bezüglich ihrer individuellen Eigenschaften insbesondere in der Anfälligkeit (Frailty) gegenüber Krankheiten und dem Ereignis Tod unterscheiden.

Die Frailty-Varianz kann somit auch als Mixture-Variable durch Mischverteilungen modelliert werden, um somit diese Variabilität innerhalb der Anfälligkeitsmerkmale zu beachten.

Dabei erweist sich die Anwendung der Gamma-Verteilung für die Modellierung der Frailty-Varianz als sehr vielseitig (Vaupel et al., 1979; Lancaster, 1979; dos Santos et al., 1995, Hougaard, 2000).

Die Schätzungen mit Gamma-Frailty-Modellen lassen sich methodisch mit relativ geringem rechnerischen Aufwand realisieren, und eine direkte Beziehung zum unbedingten Survival lässt sich auch herstellen. Die Gamma-Verteilung ist auch deshalb bei vielen Anwendungen beliebt, weil sie als eine der flexibelsten statistischen Verteilungsfunktionen nahezu alle möglichen Verteilungsfunktionen abbilden kann.

Die Anpassung an eine bestimmte Verteilungsfunktion kann anhand der Variation des Form-Parameters k bestimmt werden.

Wenn $k = 1$ ist, dann folgt die Verteilungsfunktion der Exponentialverteilung. Ist der Wert für $k > 1$, dann nähert sich diese Funktion zunehmend der Normalverteilung an.

Angesichts dieser Vorteile muss aber auch erwähnt werden, dass keine biologischen Erklärungsfaktoren bestehen, welche in irgendeiner Form rechtfertigen, warum die Gamma-Verteilung den anderen statistischen Verteilungsfunktionen vorzuziehen sei (Yashin et al., 2008).

Der Frailty für die Überlebenden zu jedem Zeitpunkt t ist gammaverteilt mit dem gleichen Niveauwert des Formenparameters k für alle Individuen unter Risikobeginn bzw. zu Beginn des Follow-ups. Der Wert des Skalenparameters λ berechnet sich aus $\lambda + \Lambda_0(t)$, mit $\Lambda_0(t)$ als kumulative Baselinefunktion (vgl. Wienke, 2006, S. 31).

Die Frailty der Verstorbenen für jedes Alter t ist ebenfalls gammaverteilt und zunächst identisch mit dem Wert der Überlebenden $\lambda + \Lambda_0(t)$ zum Zeitpunkt t , allerdings verändert sich das Niveau des Formenparameters um $k + 1$.

Die Frailty-Neigungen für die Überlebenden $\frac{\lambda}{\lambda + \Lambda(t)}$ und für die Verstorbenen

$\frac{\lambda + 1}{\lambda + \Lambda(t)}$ zeigen die unterschiedlich starken Wirkungen der Selektions-

mechanismen für die so genannten Hochrisiko-Fälle. Nur unter der Bedingung einer parametrischen Restriktion von $\lambda = k$ für die Gamma-Verteilung

folgt $E(Z) = 1$ und $\sigma^2 = \frac{1}{\lambda}$, die Varianz für die Frailty-Variable.

Der unbedingte Population-Hazard und die Survivalfunktion werden definiert mit:

$$\lambda(t) = \frac{\lambda_0(t)}{1 + \sigma^2 \Lambda_0(t)} \quad \text{und} \quad S(t) = (1 + \sigma^2 \Lambda_0(t))^{-\frac{1}{\sigma^2}} \quad (4.3)$$

Dieses Modell wurde erstmalig von Vaupel et al. (1979) eingeführt.

4.5.2 Inverse Gauss-Verteilung

Als eine Alternative für die statistische Modellierung der Frailty-Varianz durch eine Gamma-Verteilung wurde von Hougaard (1984) die inverse Gauss-Verteilung eingeführt.

Wie bei der Gamma-Verteilung muss die Bedingung, dass Z invers normalverteilt mit einem $E(Z) = 1$ und einer Varianz σ^2 ist, gegeben sein.

Die Gleichung für den unbedingten Hazard unterscheidet sich im Vergleich zum Gamma-Verteilungs-Ansatz nur im Nenner, indem die verdoppelte Varianz unter Bildung eines Kehrwerts aus der Quadrierung des Nenners vorgenommen wird.

$$\lambda(t) = \frac{\lambda_0(t)}{(1 + 2\sigma^2 \Lambda_0(t))^{1/2}} \quad (4.4)$$

Die Survivalfunktion für die inverse Gauss-Verteilung wird dabei wie folgt definiert als:

$$S(t) = e^{\frac{1}{\sigma^2}(1 - \sqrt{1 + 2\sigma^2 \Lambda_0(t)})} \quad (4.5)$$

4.6. Zielsetzung aller Frailty-Ansätze

Alle dargestellten Ansätze versuchen gemäß ihren Modellannahmen die Existenz von unbeobachteter Heterogenität nachzuweisen. Dabei wird durch die Formulierung einer Zufallsvariable Z versucht, mögliche Effekte von unbeobachteter Heterogenität zu berücksichtigen, die entstehen können, wenn bei der Modellbildung bestimmte erklärende Variablen nicht beachtet werden, weil kein sichtbar feststellbarer direkter Zusammenhang mit den Regressoren X vermutet wird.

Die Modelle basieren auf dem Proportional-Hazards-Ansatz von Cox, mit dem Unterschied, dass eine Spezifizierung der Baselinefunktion parametrisch durch den Frailty-Term Z erfolgt.

$$\lambda(t, X, Z) = Z \lambda_0(t) e^{\sum_{i=1}^k \beta_i X_i} \quad (4.6)$$

Die individuelle Frailty-Neigung Z ist eine Zufallsvariable mit dem Erwartungswert von 1 mit einer parametrischen Verteilungsfunktion, die unabhängig von der Verteilungsfunktion der beobachteten Kovariablen ist.

Die Modellschätzungen erfolgen nach dem Maximum-Likelihood-Prinzip, wenn dem Baseline-Hazard $\lambda_0(t)$ eine parametrische Verteilungsfunktion und der Zufallsvariablen Z (Frailty) auch eine bestimmte statistische Verteilung zugeordnet wird.

Das Problem der Identifikation von unbeobachteter Heterogenität:

Die Proportional-Hazard-Bedingung für eine proportionale Beziehung zwischen Sterblichkeit und der Lebenserwartung (gilt nur in homogenen Bevölkerungen) (Yashin et al., 2008) wird durch die unbeobachtete Heterogenität aufgehoben.

Eine Veränderung des vorausgesetzten proportionalen Verhältnisses zu einem nichtproportionalen Zusammenhang kann ausgelöst werden, sobald ein scheinbar wichtiger Regressor aus dem Schätzmodell herausgenommen wird.

Sobald die Existenz von unbeobachteter Heterogenität vermutet wird, muss auch eine Spezifizierung der Baselinefunktion vorgenommen werden, da nun in diesem konkreten Fall kein konstanter Hazard mehr unterstellt werden kann. Die Parametrisierung der Frailty-Varianz und des individuellen Hazards kann durch jegliche Form der statistischen Verteilungsfunktionen erfolgen, wobei dies von der wissenschaftlichen Fragestellung und der Skalierung der verwendeten Variablen abhängig ist.

4.7 Vollparametrische Frailty-Modelle

Der wohl am meisten etablierte vollparametrische Frailty-Ansatz ist das Gamma-Gompertz-Modell, bekannt auch als Perks-Modell, welches in einer aktuellen Veröffentlichung (Vaupel et al., 2009) anhand der Human Mortality Database bereits überprüft wurde.

Der formale Ausdruck in 5.4 kann vollparametrisch für das Gamma-Gompertz-Modell beschrieben werden.

Perks-Modell (P) (vgl. Vaupel et al. 2009, S. 3)

$$\lambda(t) = \frac{ae^{bt}}{1 + \frac{\sigma^2_z a}{b}(e^{bt} - 1)} \quad (4.7)$$

wobei $\lambda_0(t) = ae^{bt}$ und $\Lambda_0(t) = \frac{a}{b}(e^{bt} - 1)$ gilt.

Das Perks-Modell lässt sich modifiziert auch in Form eines Gamma-Gompertz-Makeham-Ansatzes (Modified Perks MP) darstellen (vgl. Butt und Habermann, 2004, S. 178).

$$\lambda(t) = \alpha + \frac{ae^{bt}}{1 + \frac{\sigma^2_z a}{b}(e^{bt} - 1)} \quad (4.8)$$

Alternativ lässt sich für die Modellierung der Frailty-Varianz in Form einer Gamma-Verteilung relativ einfach durch eine Inverse-Gauss-Verteilung parametrisch bestimmen:

$$\lambda(t) = \frac{e^{-d+pt'}}{\sqrt{1 + e^{-b+pt'}}} \quad (4.9)$$

Durch Berücksichtigung des Faktors t' können Altersangaben adjustiert werden, sodass ein Mindestalter bestimmt wird und damit die parametrischen Schätzungen wesentlich genauer werden. Diese Maßnahme erweist sich auch für die Fragestellung dieser wissenschaftlichen Arbeit als sehr sinnvoll, da im Datensatz einige Personen enthalten sein können, die ihre Familienplanung scheinbar schon abgeschlossen hatten und eine andrologische Untersuchung aus anderen Gründen haben durchführen lassen.

5. Daten- und Projektbeschreibung

5.1. Forschungsleitende Hypothesen

Die statistische Analyse erfolgt unter Berücksichtigung folgender Fragestellungen:

1. Unterscheiden sich subfertile und fertile Fälle bezüglich ihres Survivalverlaufs?
2. Sind potentielle Sterblichkeitsunterschiede real oder resultieren sie aufgrund von unbeobachteter Heterogenität?
3. Wird die Proportional-Hazards-Annahme verletzt?
4. Ist die Existenz von unbeobachteter Heterogenität nachweisbar; wie stark ist ihre Intensität?
5. Ergeben sich Unterschiede bezüglich der Parameterschätzungen im Vergleich von Unshared-Frailty-Modellen.
6. Welches statistische Modell liefert die adäquateste Anpassung für die Spermioigrammdaten?

5.2 Projektbeschreibung

Die vorliegende Arbeit ist Teil eines seit 2003 laufenden Forschungsprojekts zur Analyse der Sterblichkeitsunterschiede von Männern mit normalen und subnormalen Spermienkonzentrationen. Für die bisherige Analyse wurden 1283 ehemalige Patienten berücksichtigt, von denen für 601 Fälle der Vitalstatus überprüft werden konnte (Groos, 2005).

In dieser Studie erfolgte eine Erweiterung der Studienpopulation um mehr als 1.000 Personen und eine Aktualisierung der Angaben zum Vitalstatus für einen späteren Stichtag (31.12.2006).

Parallel werden zwei weitere wissenschaftliche Arbeiten – (Seydel, 2010): Überlebenszeit und Todesursachen von fertilen und subfertilen Männern sowie Belting (2011): Lebenslaufmorbidity und Reproduktionsbiographie bei fertilen und subfertilen Männern – zu jeweils unterschiedlichen Fragestellungen verfasst, die sich ebenfalls auf diese Datengrundlage beziehen.

5.3 Datengrundlage

Als Datengrundlage dienen Patientenakten der Fertilitätssprechstunde der andrologischen Abteilung des Universitätsklinikums Marburg (Hessen) aus den Untersuchungsjahren von 1949 bis 1998. In die Stichprobe wurden hauptsächlich jene Patienten aufgenommen, die sich aufgrund eines unerfüllten Kinderwunsches innerhalb des benannten Zeitraums einer Untersuchung ihrer SpermioGrammparameter unterzogen haben.

Ein weiteres wichtiges Einschlusskriterium für die Probanden war die Bedingung, vor dem 01.01.1942 geboren zu sein, damit zum Stichtag der Überprüfung des Vitalstatus am 31.12.2006 sicher war, dass alle Patienten mindestens das 65. Lebensjahr vollendet hatten.

Die wenigen Patienten nichtdeutscher Herkunft bzw. mit Migrationshintergrund wurden für die Untersuchungspopulation nicht berücksichtigt, um sicherzustellen, dass die Untersuchungspersonen bezüglich ihrer Lebensstile und ihres Gesundheitszustandes möglichst homogen seien.

5.4 Patientenakten

Die für die Studie verwendeten Patientenakten haben sich während der gesamten fast fünfzigjährigen Untersuchungsphase erheblich gewandelt.

Insgesamt wurden sechs verschiedene Formbögen für die Spermioграмmuntersuchung verwendet, die sich insbesondere in der Dokumentation der Befunde und Anamnese unterscheiden.

Ältere Befunddokumentationen, überwiegend aus den 40er und 50er Jahren, weisen viel häufiger detaillierte Angaben zum persönlichen Sexualverhalten, dem Fertilitätsstatus der Ehepartnerin, verhaltensspezifische Besonderheiten wie regelmäßigen Alkohol- und Zigarettenkonsum oder psychische Erkrankungen aufgrund von Traumata auf. Mit den aktuelleren Erhebungsbögen werden die Anamnese und die organischen Befunde nur in kurzer Form erhoben und lassen daher vermuten, dass nicht alle relevanten Erkrankungen erfasst sind.

Vielmehr erfolgt eine stärkere Konzentration auf die primäre Auszählung der Spermindichten und die Bestimmung anderer Fertilitätsparameter wie Morphologie und Motilität.

5.5 Vollständigkeit und Konsistenz der Patientenakten

Für die anstehende statistische Analyse ist zu beachten, dass nicht alle Spermio-gramm-Formblätter für alle Patienten in gleicher Weise vollständig ausgefüllt wurden, sodass einzelne Variablen mehr oder weniger mit Fehlwerten behaftet sind. Nur wenige Angaben erweisen sich als einigermaßen lückenlos dokumentiert, wie die demographischen Angaben zu Geburtsdatum, Alter, Untersuchungsjahr sowie die Spermio-gramm-parameter wie Spermindichte, pH-Wert, Morphologie und Motilität.

Für einige Patienten wurden ebenfalls Begleiterkrankungen wie Mumps, Orchitis, Kryptorchismus, Varikozele sowie operative Eingriffe wie beispielsweise eine Vasektomie aufgenommen. Bezüglich der Komorbiditäten gilt zu entscheiden, ob bestimmte Fälle aufgrund einer möglicher Erkrankung aus der Studienpopulation ausgeschlossen werden sollten. Da nicht sicher ist, ob die Dokumentation von fertilitätsbeeinträchtigenden Begleitfaktoren nachweislich vollständig ist, kann dieses einen Ausschluss von einem gut und vollständig dokumentierten Fall zur Folge haben, dass gleich gelegene Fälle weiterhin in der Untersuchungspopulation bestehen.

Dadurch würde sich eine zusätzliche Form von Ausselektion ergeben, die sich ohnehin schon aufgrund von fehlender Vollständigkeit bei einigen Variablen ergibt. Hinzu kommt, dass möglicherweise einige Krankheiten zeitlich zu weit zurückliegen, sodass sich einzelne Personen, bedingt durch den Erinnerungsbias, nicht mehr erinnern können.

In diesem Zusammenhang stellt sich auch die Frage, welche der drei vorhandenen Spermienparameter – Spermienkonzentration, Morphologie und Motilität – die Fertilitätseigenschaften der untersuchten Patienten am besten wiedergeben. Eine rein objektive und damit auch konsistente Beurteilung über den gesamten Beobachtungszeitraum ist freilich nur bei der Auszählungsmethode der Spermienkonzentration in der Zählkammer gewährleistet, während die Bestimmung der beiden anderen Parameter doch eher überwiegend subjektiv erfolgt und eher von der fachlichen Qualifikation und den Erfahrungen des Untersuchungspersonals abhängt. Daher wird für die statistische Analyse nur die Spermienkonzentration als einziger Fertilitätsparameter verwendet. Wenn für die Patienten mehrere in zeitlichen Abständen folgende Spermio-grammuntersuchungen vorlagen, wurde nur die jeweils erste für die Auswertung berücksichtigt.

Die Abgrenzung der einzelnen Fertilitätszustanden erfolgt gemäß der Klassifikation der WHO (2010) mit 0 Mio./ml für azoosperm, von weniger als 15 Mio./ml für oligozoosperm und mit mehr als 15 Mio./ml.

5.6 Recherche und Überprüfung des Vitalstatus

Eine erste Überprüfung des Vitalstatus für die ehemaligen Patienten erfolgte bereits am 31.12.2002. Diese Recherche bezog sich aber nur auf eine kleinere Stichprobe von 1283 Personen und schloss nur Patienten ein, die vor dem 01.01.1938 geboren wurden. In einer Kooperation mit dem Einwohnermeldeamt der Stadt Marburg konnte geprüft werden, ob die ehemaligen Patienten dort mit einem Wohnsitz gemeldet waren. Die Identifizierung der einzelnen Fälle erfolgte über die persönlichen Angaben zum Vornamen, Familiennamen sowie dem Geburtsdatum, welche auf der Krankenakte vermerkt waren. Da die Registrierungen der Meldeinformationen seit den 1980er Jahren nicht mehr über ein papierschriftliches Karteisystem erfolgen, sondern in elektronischer Form als EDV-Dateien vorliegen, war Suche nach dem aktuellen Wohnsitz erheblich vereinfacht.

Für einige Fälle, die überwiegend vor diesem Zeitpunkt als Verstorbene bzw. Verzogene ermittelt wurden, konnte der Vitalstatus nicht über den PC ermittelt werden, sodass zusätzlich noch eine weitere Recherche im Archiv des Einwohnermeldeamtes erforderlich war. Konnten die Patienten über das Meldesystem gefunden werden und waren bis zum diesem Zeitpunkt noch am Leben, wurden der gegenwärtige Vitalstatus und die aktuelle Meldeadresse vermerkt.

Sobald aber eine Person verstorben war, wurden alle relevanten Angaben zum Sterbedatum und Sterbeort festgehalten.

Bei Fortzügen wurde die aktuelle Adresse über das Einwohnermeldeamt Marburg eingeholt und eine entsprechende Anfrage bei der jeweiligen zuständigen Gemeinde getätigt.

Etwas aufwendiger gestaltete sich die Suchprozedur, wenn die Person mehrfach umgezogen war, da mehrere Einwohnermeldeämter angeschrieben werden mussten. Ein Teil an Informationen über den Vitalstatus konnte auch über das hessische Zentralarchiv einer gesetzlichen Krankenkasse ermittelt werden.

Die zweite Recherche bezieht sich dagegen auf den Stichtag 31.12.2006 und knüpft direkt an die erste Suche an. Die bereits vorhandenen Informationen zum Vitalstatus wurden übernommen und ausreichend dokumentiert.

So konnten insgesamt 52,5% der Fälle direkt über das Einwohnermeldeamt bestimmt werden, die Vitaldaten von 8,4% aller gefundenen Patienten wurden über das hessische Zentralarchiv der gesetzlichen Krankenversicherung gewonnen, 39,2% waren sowohl über die Krankenkasse als auch über das Einwohnermeldeamt auffindbar.

5.7 Die aktuelle Untersuchungspopulation

Die Erweiterung der Studienpopulation umfasst eine Stichprobe von insgesamt 2296 Fällen. Für 1408 ehemalige Patienten konnte der Vitalstatus für den Stichtag 31.12.2006 überprüft werden, die übrigen 888 Fälle gelten als Lost-to-Follow-up, konnten entsprechend nicht nachverfolgt werden.

Für die Survivalanalyse problematisch erweist sich der Zustand eines relativ hohen Anteils von nichtverstorbenen Fällen (974 gegenüber 434 Ereignissen). Ebenso besteht eine Varianz der betrachteten Fälle bezüglich des Geburtsjahres, des Untersuchungsjahres sowie des Untersuchungsalters.

Personen, die während des Ersten Weltkrieges geboren wurden, hatten bedingt eine insgesamt zu beobachtende Sterblichkeitsverbesserung, möglicherweise eine höhere allgemeine Sterbeneigung als Personen, die erst in den dreißiger oder vierziger Jahren geboren wurden.

Ebenso war festzustellen, dass innerhalb der letzten 50 Jahre sechs verschiedene Versionen der Spermogramm-Formblätter genutzt worden waren.

Es ist somit nicht auszuschließen, dass die Erhebungsmethoden bzw. die Genauigkeit bei der Bestimmung der Spermienparameter (Spermiedichte, Morphologie und Motilität) gemäß dem technologischen Fortschritt variieren kann. Patienten suchen aus unterschiedlichen persönlichen Gründen die andrologische Abteilung auf.

Es ist aber auch durchaus vorstellbar, dass je nach Lebensalter die individuellen Motive bezüglich der Spermogrammuntersuchung unterschiedlich sein könnten. Männer im jungen Erwachsenenalter zwischen 30 und 40 Jahren würden sich wohl eher aufgrund eines noch nicht erfüllten Kinderwunsches untersuchen lassen.

Ältere Männer etwa ab dem 50. Lebensjahr haben wohl möglicherweise häufiger mit der Familienplanung bereits abgeschlossen und lassen sich häufiger aufgrund von Erektionsstörungen oder anderen sexuellen Störungen, die nicht unmittelbar im Zusammenhang der Reproduktion stehen, behandeln. Andererseits könnten diese Männer aber auch von einer „sekundären Infertilität“ betroffen sein, was bedeutet, dass sie bereits in einer früheren Partnerschaft erfolgreich Kinder gezeugt haben, nun aber in einer neuen Beziehung mit einer jüngeren Partnerin leben, bei der eine ungewollte Kinderlosigkeit festgestellt wird.

In der Arbeit von Seydel (2010) wurden für die statistische Analyse zunächst nur 1202 normo- und oligozoosperme Fälle mit bekanntem Vitalstatus berücksichtigt. Ein Ausschluss der azoospermen Fälle erfolgte mit der Begründung, dass eine Azoospermie eher auf eine Verschlussstörung zurückgeführt werden kann.

Eine Beeinträchtigung der Spermatogenese muss daher nicht zwangsläufig folgen. Eine Azoospermie kann durchaus viele Ursachen haben, wie auch in der Studie von Jensen et al. (2009) richtig argumentiert wird. Dennoch kann eine Nichtberücksichtigung von Fällen mit bekanntem Vitalstatus nur dann erfolgen, wenn die Stichprobe wie bei Jensen et al. (2009) sehr groß ist.

Zudem ist eine gewisse Heterogenität der möglichen Ursachen auch bei oligozoospermen Fällen zu beobachten.

Aus diesem Grund werden für die nachfolgenden Untersuchungen auch die azoospermen Fälle berücksichtigt.

5.8 Definition von Lebenszeit und Beobachtungszeit

Die Lebenszeit für die untersuchten Fälle wird definiert als die gesamte Lebensdauer ab dem jeweiligen Geburtsdatum bis zum Sterbedatum bzw. Ende der Follow-up- Periode (31.12.2006). Eine solche Betrachtung erfolgte bereits in der Arbeit von Groos, 2005, allerdings ohne eine korrekte Altersadjustierung zu berücksichtigen. Dabei ist zu beachten, dass die betrachtete Beobachtungszeit durch Linkstrunkierungen beeinflusst wird. Die Beobachtungszeit beginnt dabei schon bereits mit dem Zeitpunkt der Geburt, weil ab da an jeden einzelne Person dem Risiko ausgesetzt ist. Im Idealfall wäre das Risiko für alle Personen gleich.

Da der Beobachtungsbeginn innerhalb der Studienpopulation sehr variiert, kann eine solche Betrachtung so ohne weiteres nicht vorgenommen werden. Die geschätzte Lebensdauer eines jeden Individuums ist somit von der Länge des Beobachtungszeitraums abhängig und wird daher von dem Geburtstag beeinflusst. Dabei haben einige Probanden bereits eine gewisse Dauer an Überlebenszeit verbracht, bevor sie tatsächlich unter Beobachtung standen. Sie konnten auch nur deshalb beobachtet werden, da sie bis zu diesem Zeitpunkt noch nicht verstorben waren. Gelöst wird dieses Problem durch eine Modellierung des delayed entries die Berücksichtigung verschiedener individueller Zeitpunkte t_{0i} bei Studieneintritt.

Dabei wird wie folgt vorgegangen: Es wird zunächst der individuelle Zeitpunkt des Beobachtungsbeginns t_{0i} für jede Untersuchungsperson (die zeitliche Differenz in Tagen aus dem Datum der Untersuchung und Geburtstag) gemessen und dann die Beobachtungszeit t_i , die Zeitdauer (Überlebenszeit) als zeitliche Differenz in Tagen vom Todes- bzw. Zensierungsdatum bis Geburtstag gebildet. Für die Likelihood-Funktion (vgl. Cleves et al., 2008) gilt dann folgende Bedingung:

$$L_i(\beta|t_i, x_i) = \{S(t_i|x_i, \beta) / S(t_{0i}|x_i, \beta)\} \lambda(t_i|x_i, \beta).$$

Entscheidend ist dabei $S(t_i|x_i, \beta) / S(t_{0i}|x_i, \beta)$ das Verhältnis aus dem Survival bis zum Zeitpunkt t_i und dem Survival zum Zeitpunkt t_{0i} . Zwei weitere Bedingungen für die Definition von Überlebenszeit mit delayed entry werden vorausgesetzt, die unbedingt erfüllt werden müssen (Hosmer, et al. 2008):

1. Für jede Beobachtung muss anhand des Studiendesigns ein Mindestwert an Überlebenszeit bestimmt werden können, der für alle Probanden entweder gleich oder unterschiedlich ist.
2. Der Zeitpunkt bei Studieneintritt muss für jede Beobachtung bekannt sein.

Bei der Berücksichtigung von delayed entry ist es mitunter notwendig einzelne Untersuchungspersonen von der Analyse auszuschließen, weil die aufgeführten Bedingungen nicht erfüllen werden können.

Dieses ist immer dann der Fall, wenn einzelne Untersuchungspersonen erst zu einem relativ späten Zeitpunkt in die Studie aufgenommen wurden und bis zu diesem benannten Zeitpunkt bereits alle Sterbeereignisse stattgefunden haben.

Ebenso ist vorstellbar, dass ein später Beitritt eines Individuums zur Untersuchungspopulation, dazu führen kann, dass die Zeit unter Beobachtung dann relativ kurz ist und ein Ausscheiden wohl eher altersbedingt ist und dann nicht auf den Risikofaktor zurückgeführt werden kann.

Werden diese Probleme nicht ausreichend beachtet, kann dieses im Extremfall zu starken Verzerrungen der Schätzergebnisse führen.

Eine ausführliche Darstellung der methodischen Vorgehensweise bei Linkstrunkierung bzw. von delayed entries in der Überlebenszeitanalyse erfolgt u.a. in Pan und Chappell, 2002, Matsuura und Eguchi, 2005, Cleves et al., 2008, Hosmer et al., 2008, sowie in Anzuers-Cabrera und Hutton, 2010.

Für die folgende Auswertung wird nur die Beobachtungszeit ab Geburtstag bis zum Todes- bzw. Zensierungstag genutzt, weil nur in diesem Fall eine sinnvolle Behandlung der Linkstrunkierungen erzielt wird.

Durch diese Maßnahmen werden die unterschiedlichen Untersuchungsalter der Personen adjustiert, um damit die unterschiedlichen individuellen Lebenszeiten miteinander vergleichen zu können. Das hat ebenfalls zur Folge, dass die Variable Untersuchungsalter im Regressionsmodell nicht mehr als separate Kovariable aufgeführt wird.

Da nicht für alle untersuchten Fälle vollständige Angaben zum Datum der Untersuchung vorhanden waren bzw. nicht immer das exakte Alter bei der Untersuchung ermittelt werden konnte, wird sich die anschließende Datenanalyse auf 974 normozoosperme, 225 oligozoosperme und 205 azoosperme Fälle (insgesamt 1404 Fälle) beschränken.

Die Zahl der Sterbeereignisse beträgt 433. Unter der Berücksichtigung des delay-entry mussten zusätzlich noch zwei weitere Fälle ausgeschlossen werden, weil wie bereits erwähnt die aufgeführten Bedingungen bei einzelnen Untersuchungspersonen nicht erfüllt werden konnten.

5.9 Aufstellung der verwendeten Kovariablen

Für die semi- und vollparametrischen Schätzverfahren werden neben dem Fertilitätsstatus noch weitere Erklärungsfaktoren wie das Geburtsjahr und der Wohnort berücksichtigt.

Tabelle 5: Aufstellung aller verwendeter Variablen

Variable	Variablentyp	Ausprägung	Einheit
Spermienkonzentration	numerisch	0- 550	Mio./ml
Fertilitätsstatus (subfertil, fertil) WHO (2010)	klassiert	subfertil = 0 bis unter 15 fertil = 15 und mehr	Mio./ml Mio./ml
Fertilitätsstatus (azoosperm, oligozoosperm, normozoosperm) WHO (2010)	klassiert	azoosperm = 0 oligozoosperm = größer 0 bis unter 15 normozoosperm = 15 und mehr	Mio./ml Mio./ml Mio./ml
Geburtsjahr	numerisch		Jahre
Wohnort	klassiert	Stadt Marburg Landkreis Marburg Außerhalb des Landkreises Marburg	- - -

Für die statistische Analyse werden für die Erfassung der Fertilitätskomponente die Spermienkonzentration als numerische Größe verwendet.

Ebenso erfolgt die Berücksichtigung des Fertilitätsstatus gemäß der Klassifikation der WHO 2010, dadurch ist möglich die Hazard-Ratios für die Vergleichsgruppen zu bestimmen.

Im direkten Vergleich der Hazard-Ratios von fertilen und subfertilen Fällen, werde jene Patienten als Referenzgruppe definiert, die normale Spermienparameter auf gewiesen haben. Gleiches gilt auch für die Vergleiche der

Hazardneigung von azoospermen, oligozoospermen und normzoospermen Fällen, bei denen letztgenannte die Referenzgruppe bilden.

Als weitere Kovariable wird das Geburtsjahr wird als mögliche Erklärungsgröße für Sterblichkeitsunterschiede herangezogen, da mitunter die Mortalität auch zwischen den Geburtsjahrgängen variieren kann.

Die Kovariable Wohnort soll darüber Auskunft geben, ob räumliche Aspekte die Sterblichkeit beeinflussen. Dabei könnte man vermuten, dass die Untersuchungspersonen, welche von außerhalb des Landkreises Marburg stammen möglicherweise schwerwiegende Fertilitätsprobleme haben und diese daher in einer spezialisierten Einrichtung behandeln lassen möchten.

5.10 Definition der verwendeten statistischen Parameter und Prüfgrößen

Für die folgende statistische Analyse der Regressionsmodelle werden betrachtet:

1. Der Regressionskoeffizient (Coef.) beschreibt im Rahmen der Regressionsanalyse welchen Beitrag die unabhängige Variable für die Prognose der abhängigen Variable liefert.
2. Der Standardfehler (SE) beschreibt die durchschnittliche Abweichung des geschätzten Parameterwertes vom wahren Parameterwert. Dabei gilt, je geringer der Wert des Standardfehlers ist, desto genauer kann der unbekannte Schätzparameter mit der Schätzfunktion ermittelt werden.
3. Das Hazard-Ratio (HR) beschreibt den statistischen Einfluss einer unabhängigen Variable auf das Hazardrisiko. Der Likelihood-Ratio-Test ist eine Testprozedur zum Vergleich von Modellen auf der Grundlage des Maximum-Likelihood-Schätzers.

Dabei werden jeweils zwei Modelle verglichen: Ein Ausgangsmodell, das mehrere Modellparameter enthält. Dieses Modell wird auch als unrestringiertes Modell bezeichnet. Das Vergleichsmodell, welches das restringierte Modell darstellt enthält dagegen eine oder mehrere Parameterrestriktionen. Mit dem Likelihood-Ratio-Test wird geprüft, ob das unrestringierte Modell eine signifikante bessere Modellanpassung liefert als das restringierte Modell.

Die Testprozedur folgt dabei der Chi-Quadrat-Verteilung unter der Berücksichtigung der Zahl der Freiheitsgrade (df), welche der Anzahl der Restriktionen entspricht.

4. Der Koeffizient θ beschreibt die Prüfgröße für die Frailty-Varianz, welche ebenfalls auf dem Prinzip des Likelihood-Ratio-Tests aufbaut.
5. Die statistische Signifikanz wird durch Irrtumswahrscheinlichkeit p mit * bei $p \leq 0,05$ und ** bei $p \leq 0,01$ gekennzeichnet.

5.11 Statistik-Software

Alle analytischen Verfahren wie deskriptive Statistik, nichtparametrische, semi- und vollparametrische Modellrechnungen wurden mit der Statistik-Software STATA 10 SE durchgeführt. Auch die graphischen Darstellungen erfolgten mit diesem Programm. Weitere Informationen zur Software sind unter <http://www.stata-press.com> abrufbar.

6. Ergebnisse

6.1 Zusammensetzung des verwendeten Datensatzes

Es folgt eine Beschreibung von einigen relevanten Variablen, die bei den folgenden Parameterschätzungen berücksichtigt werden.

Die Darstellung insbesondere der Lagemaße und relativen Häufigkeiten orientiert sich an der Vorgehensweise in der Arbeit von Groos (2005).

Dadurch ist es möglich, auf eventuelle Veränderungen in der Zusammensetzung der aktuellen Fälle hinzuweisen, welche eventuell entstehen können, sobald eine Erweiterung des ursprünglichen Datensatzes mit z.B. deutlich jüngeren Untersuchungspersonen erfolgt. Sollten sich durch die Imputation weiterer Fälle die Kompositionen bezüglich der Anteile der Sterbefälle, der Spermienkonzentration oder das Untersuchungsalter verändern, ist ein eventueller Vergleich der Ergebnisse nicht mehr möglich.

Tabelle 6: Vitaldaten der Untersuchungspopulation

n	1408
Geburtsjahr	1892-1941
mean	1932,99 (SD 7,2)
1. Quartil (25 %)	1929
2. Quartil (50 %)	1935
3. Quartil (75 %)	1939
Vitalstatus am 31.12.2006	
lebend	973 (69,15 %)
gestorben	434 (30,85 %)
Sterbejahr	1952-2006
mean	1995, 97 (SD 9,07)
1. Quartil (25 %)	1991
2. Quartil (50 %)	1998
3. Quartil (75 %)	2003

Die Spannweite für die Variable Geburtsjahr erstreckt sich von Ende des 19. Jahrhunderts bis zum Jahr 1941, d.h., der älteste Patient war zu diesem Zeitpunkt schon mindestens 49 Jahre alt, bevor der jüngste geboren wurde. Knapp ein Drittel aller betrachteten Fälle ist bis zum Stichtag 31.12.2006 verstorben, diese Zahl entspricht ungefähr dem Wert, welcher in der Arbeit von Groos (2005) ermittelt wurde.

Auffällig ist auch eine Anhäufung von Fällen, die nach dem Jahr 1937 geboren wurden; Dieses sind dann entsprechend die Fälle, welche aufgrund der Fortsetzung der Follow-up-Periode erweitert wurden.

Der relative Anteil (34,2%) von Patienten, die nach 1937 geboren wurden, hat sich dementsprechend erhöht und könnte daher eine potentielle Heterogenitätswirkung bezüglich der Sterbeneigung zur Folge haben, da später geborene Personen bekanntlich ein geringes Sterberisiko haben könnten.

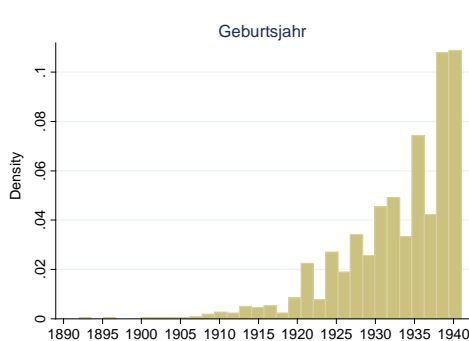


Abbildung 3: Histogramm Geburtsjahr der Fälle

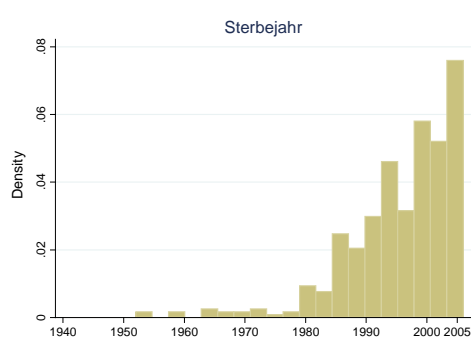


Abbildung 4: Histogramm Sterbejahr der Patienten

Knapp 80% der Sterbefälle ereigneten sich nach 1990, wovon ein relativer Anteil von 34,6% erst nach dem Jahr 2000 zu verzeichnen war.

Tabelle 7: Individualdaten und Fertilitätsklassifikation aus den Patientenakten

Jahr der Untersuchung (n=1405)	1949-1994
mean	1967,41 (SD 7,7)
1. Quartil (25 %)	1962
2. Quartil (50 %)	1967
3. Quartil (75 %)	1972
Untersuchungsalter (n=1404)	18,9-67,1
mean	34,67 (7,32)
1. Quartil (25 %)	29,7
2. Quartil (50 %)	33,0
3. Quartil (75 %)	37,9
Spermogramm (Mio./ml)	
Klassifikation der WHO (n=1408)	
subfertil	433 (30,7 %)
fertil	975 (69,3 %)
azoosperm	206 (14,63 %)
oligozoosperm	227 (16,12 %)
normozoosperm	975 (69,24 %)
oligozoosperm (n=227)	0,2-14
mean	6,39 (SD 4,26)
1. Quartil (25 %)	2
2. Quartil (50 %)	6,25
3. Quartil (75 %)	10
normozoosperm (n=975)	15-550
mean	92,34 (SD 72,20)
1. Quartil (25 %)	41
2. Quartil (50 %)	75
3. Quartil (75 %)	121

Die Angaben der Tabelle verweisen auf eine sehr große Spannweite bezüglich der Untersuchungsperiode von knapp 50 Jahren.

Daher könnten vermutlich periodische Effekte, welche z.B. bei einer Optimierung der Messmethoden für die Spermienparameter entstehen, innerhalb der Beobachtungsphase auch einen Einfluss auf die Parameterschätzung in Bezug auf die Ermittlung eines Zusammenhangs von Sterblichkeit und Fertilität haben.

Auch die Altersspanne der Patienten zum Zeitpunkt der Untersuchung variiert sehr stark, zwischen 19 und 67 Jahren, und zeigt, dass sich die Patienten womöglich aus unterschiedlichen Gründen und Motiven untersuchen lassen.

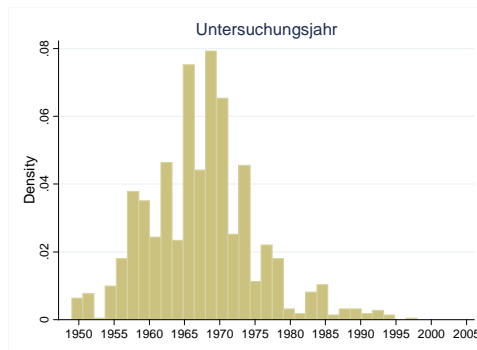


Abbildung 5: Histogramm Jahr der ersten Untersuchung

Unter Berücksichtigung des Durchschnittsalters von knapp 35 Jahren kann aber mit ziemlicher Sicherheit davon ausgegangen werden, dass sich ein sehr großer Teil der Patienten noch in der reproduktiven Phase befindet und sich daher mit dem Thema unerfüllter Kinderwunsch auseinandersetzt.

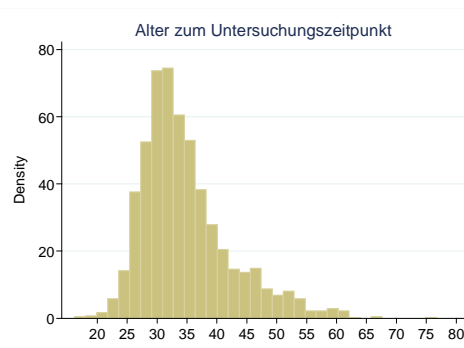


Abbildung 6: Histogramm Alter zum Zeitpunkt der ersten Untersuchung

Entscheidend für die sukzessive Analyse ist aber auch zu bestimmen, inwiefern sich die Verteilung innerhalb der Fertilitätsklassifikation, also der Differenzierung nach fertilen und subfertilen Fällen, aufgrund der Erweiterung des Datensatzes verändert hat.

So ist zu erkennen, dass mehr als zwei Drittel aller Patienten zumindest nach dem Bewertungsmaßstab der WHO (WHO, 2010) normale Spermienparameter aufwiesen, rund 16 % hatten weniger als 15 Mio./ml im Ejakulat und in 14,6% der Fälle konnten keine Spermien gefunden werden.

Mit der Anwendung der aktuellen WHO-Richtlinien aus dem Jahre 2010 und der Erweiterung des Datensatzes haben sich die Proportionen der Fertilitätsklassifikationen gegenüber den Verteilungen in der Arbeit von Groos (2005) nur geringfügig verändert.

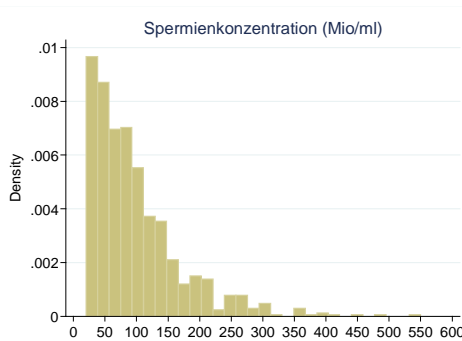


Abbildung 7: Histogramm Spermienkonzentration (Mio./ml)
für alle Patienten

Der Maximalwert für die Spermienkonzentration wird bei 550 Mio./ml erreicht, allerdings muss auch erwähnt werden, dass die mittlere Spermiedichte sehr stark von den Subfertilen, dabei insbesondere durch die azoospermen Fälle beeinflusst wird. Auf der anderen Seite ist aber auch ein weiteres Extrem zu beobachten: Etwa 6% der Patienten wiesen eine Spermienkonzentration von mehr als 200 Mio./ml auf.

Für die normozoospermen Fälle konnte eine durchschnittliche Spermienkonzentration von 92,34 Mio./ml (SD 72,20) ermittelt werden, davon hatten aber ca. 66% eine Konzentration von weniger als 100 Mio./ml.

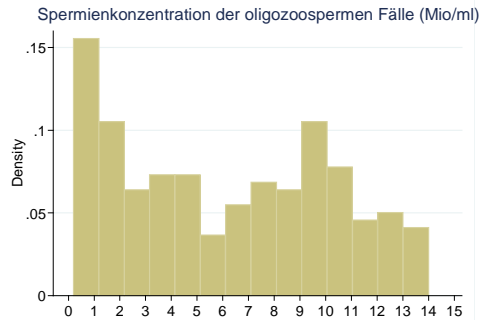


Abbildung 8: Histogramm Spermienkonzentration der oligozoospermischen Fälle

Die Verteilung der Konzentration für die oligozoospermischen Fälle scheint sich bis auf die untere bzw. die mittlere Klasse gleichmäßig zu gestalten. Für etwa ein Viertel aller oligozoospermischen Männer konnte eine Spermienkonzentration von weniger als 3 Mio./ml bestimmt werden. Das arithmetische Mittel beträgt dabei 6,39 Mio./ml (SD 4,26).

Tabelle 7: Lebensdauer der Patienten

n	1408
Lebensdauer der Verstorbenen (n=433)	40,95-93,9
mean	67,45 Jahre
1. Quartil (25 %)	61,33 Jahre
2. Quartil (50 %)	67,45 Jahre
3. Quartil (75 %)	74,3 Jahre
Lebensdauer insgesamt (n=1404)	40,95-93,9
mean	70,21 Jahre (SD 7,28)
1. Quartil (25 %)	66,55 Jahre
2. Quartil (50 %)	69,72 Jahre
3. Quartil (75 %)	74,43 Jahre

Bei einem Vergleich der mittleren Lebensdauer der Verstorbenen und aller Fälle ergibt sich für die Verstorbenen eine Verkürzung der Lebensspanne von ca. 3 Jahren; das lässt vermuten, dass sich einige Sterbefälle doch deutlich vor Beendigung der Follow-up-Periode ereignet haben.

Bei Betrachtung der Lebenszeit nach dem ersten Untersuchungstermin lässt sich eine sehr große Spannweite bezüglich der Verweildauer unter Risiko feststellen.

Drei Fälle sind bereits innerhalb eines Jahres nach dem Datum der ersten Untersuchung aus der Studienpopulation ausgeschieden, während 1,5% aller Fälle mehr als 50 Jahre unter Beobachtung standen.

6.2 Die durchschnittliche Lebensdauer nach Fertilitätsstatus

Der Vergleich der durchschnittlichen Lebensdauer nach dem Fertilitätsstatus erfolgt anhand des restriktiven Means. Damit soll überprüft werden, ob die Ausfallzeit für den jeweils zuletzt beobachtbaren Fall zensiert ist.

Bei Auftreten von Zensierungen am Ende der Beobachtungsphase wird durch den restriktiven Mean der wahre Mean unterschätzt.

In STATA können dann gegebenenfalls durch die Option des extended Medians auch für die zensierten Fälle mittlere Beobachtungszeiten mittels Approximation (ad hoc) durch eine Exponentialfunktion geschätzt werden. Diese Schätzungen sollten dann aber mit großer Vorsicht bewertet werden.

Tabelle 8: Restricted Mean für die Lebensdauer insgesamt nach Fertilitätsstatus

Fertilitätsstatus	n	restricted Mean	SD	95% CI
subfertil	430	76,99973 *	0,6167427	75,7909 - 78,2085
fertil	974	79,32977	0,5751192	78,2026 - 80,4570
Gesamt	1404	78,62227	0,4561089	77,7283 - 79,5162
Fertilitätsstatus	n	restricted Mean	SD	95% CI
azoosperm	205	77,08193 *	0,8126447	75,4892 - 78,6747
oligozoosperm	225	76,90549	0,9284123	75,0858 - 78,7251
normozoosperm	974	79,32977	0,571192	78,2026 - 80,4570
Gesamt	1404	78,32977	0,4561089	77,7283 - 79,5162

* Die letzte beobachtbare Ereigniszeit ist zensiert, der Mean wird daher unterschätzt.

Unter Berücksichtigung des restriktiven Means hätten fertile Fälle insgesamt 2,3 Jahre länger gelebt als die subfertilen Patienten. Bei den azoospermen Fällen ist für den letzten Fall die Ausfallzeit zensiert, so dass der Mean unterschätzt wird. Insgesamt scheinen aber fertile Fälle etwas mehr Beobachtungszeit in der Risikopopulation verbracht zu haben als subfertile Fälle.

Es ist dabei zu beachten, dass die aufgeführten Schätzparameter für die restriktiven Means reine Punktschätzungen sind, die sich aus der Greenwood-Formel für den Standardfehler $Var\{\hat{S}(t_p)\}$ und der Dichtefunktion $\hat{f}(t_p)$ für jedes p-te Perzentil (Klein und Moeschberger, 2002) zusammensetzen, und daher nur für das jeweilige Momentum, aber nicht für den vollständigen Survivalverlauf gültig sind.

Für den Standardfehler gilt dann: $SE\{\hat{p}\} = \frac{\sqrt{\text{Var}\{\hat{S}(t_p)\}}}{\hat{f}(t_p)}$

Eine Überprüfung auf Signifikanz der Survivalverläufe erfolgt im „overall“-Vergleich durch nichtparametrische Schätzverfahren.

6.3 Zusammenhang von Sterblichkeit und Fertilität

6.3.1 Ergebnisse der nichtparametrischen Schätzungen

Die nichtparametrischen Schätzungen von Survivalverläufen für einzelne Subgruppen im Overall-Vergleich können graphisch durch die Überlebenskurven des Kaplan-Meier-Schätzers, die kumulierte Hazardfunktion von Nelson-Aalen beschrieben werden.

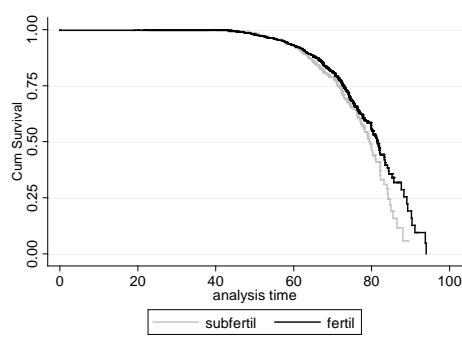
Bei vermeintlich kleinen Sterblichkeitsunterschieden ist eine rein optische Beurteilung nicht ausreichend und kann im Zweifelsfall zu falschen Interpretationen bzw. Schlussfolgerungen führen. Deshalb werden die Teststatistiken von Log rank-, Wilcoxon-, Tarone-Ware- und Peto-Peto-Prentice herangezogen, um zu überprüfen, ob sich die Überlebensfunktionen der einzelnen Vergleichsgruppen tatsächlich voneinander unterscheiden.

Der Survival-Vergleich erfolgt für die einzelnen Fälle nach Differenzierung des Fertilitätsstatus.

6.3.1.1 Vergleich der Survivalfunktionen von fertilen und subfertilen Fällen

Der Abbildung 9 ist zu entnehmen, dass die Überlebenskurve der fertilen und subfertilen Fälle bis zum Alter 60 Jahre sich kaum voneinander unterscheiden lassen. Danach entwickeln sich beide Verläufe nur marginal auseinander.

Abbildung 9: Kaplan-Meier-Survivalfunktion für subfertile und fertile Fälle



Ein signifikanter Unterschied in der Sterblichkeit von fertilen und subfertilen Fällen lässt sich somit rein visuell nicht bestimmen. Auffällig sind aber die Sprungstellen bei beiden Funktionen zum Ende der Beobachtungszeit, was sich auf den Einfluss der zensierten Fälle und die damit verbundene immer kleiner werdende Anzahl an Ereigniszeiten zurückführen lässt.

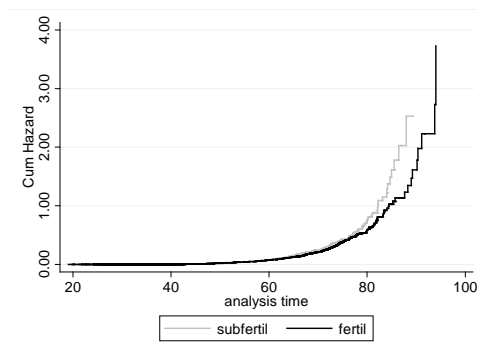
Tabelle 9: Teststatistik von Unterschieden im Survivalverlauf
für subfertile und fertile Fälle

overall	n		Log-Rank	Wilcoxon	Tarone-Ware	Peto-Peto-Prentice
Vergleich	unzensiert	zensiert	χ^2 (Pr> χ^2)	χ^2 (Pr> χ^2)	χ^2 (Pr> χ^2)	χ^2 (Pr> χ^2)
subfertil - fertil	433	971	4,13 (0,0862)	1,88 (0,1700)	2,37 (0,1239)	2,77 (0,0960)

Anhand der χ^2 -Statistik wird geprüft, ob sich die Survivalverläufe zwischen den Vergleichsgruppen unterscheiden.

Somit konnte keine der verwendeten Testprozeduren einen signifikanten Unterschied im Survivalverlauf von fertilen und subfertilen Fällen ermitteln. Der optische Eindruck aus der graphischen Darstellung (Abb.9-10) identischer Überlebensfunktionen lässt sich für beide Subgruppen bestätigen.

Abbildung 10: Kumulierte Hazards nach Nelson-Aalen für
subfertile und fertile Fälle

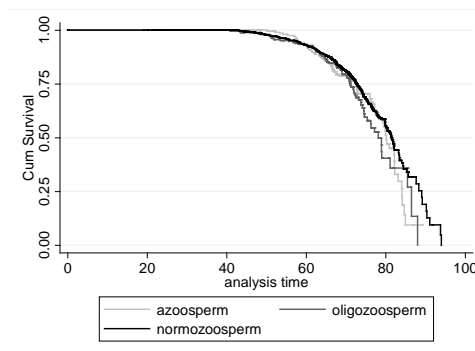


Analog zu der Darstellung der Survivalkurven nach Kaplan-Meier lassen sich auch für die kumulierten Hazards keine signifikanten Unterschiede feststellen. Insgesamt scheinen dabei die Sterblichkeitsunterschiede von subfertilen und fertilen Fällen sehr klein zu sein.

6.3.1.2 Vergleich der Survivalfunktionen von azoospermen, oligozoospermen und normozoospermen Fälle

Beim Survival-Vergleich (Abb.11 und Abb.12) für drei Untersuchungsgruppen fällt auf, dass die Sterbefunktionen von normozoospermen, oligozoospermen und azoospermen Patienten anders als erwartet doch einen sehr ähnlichen Verlauf haben.

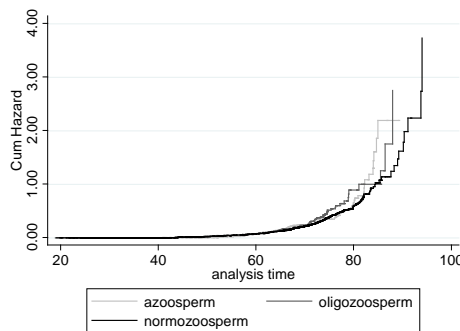
Abbildung 11: Kaplan-Meier-Survivalfunktion für azoosperme, oligozoospermen und normozoosperme Fälle



Es sind mehrfache Konvergenzen bzw. marginale Überschneidungen der Survivalfunktionen für die drei Subgruppen zu beobachten. Die Hazardverläufe scheinen sich mehrfach zu berühren.

Auffällig sind dabei auch die Sprungstellen zum Ende des Beobachtungszeitraums, was wiederum auf den Einfluss der zensierten Fälle zurückzuführen ist.

Abbildung 12: Kumulierte Hazards nach Nelson-Aalen für azoosperme, oligozoosperme und normozoosperme Fälle



Aufgrund der dargestellten Teststatistiken Tab. (10) kann im Overall-Vergleich die Nullhypothese identischer Hazards-verläufe nicht abgelehnt werden. Daher kann nicht davon ausgegangen werden, dass sich azoosperme, oligozoosperme und normozoosperme Fälle bezüglich ihrer Sterbefunktionen unterscheiden.

Tabelle 10: Teststatistik von Unterschieden im Survivalverlauf für azoosperme, oligozoosperme und normozoosperme Fälle

overall Vergleich	n		Log-Rank	Wilcoxon	Tarone-Ware	Peto-Peto-Prenice
	unzensiert	zensiert	χ^2 (Pr> χ^2)	χ^2 (Pr> χ^2)	χ^2 (Pr> χ^2)	χ^2 (Pr> χ^2)
azo-, oligozo-, und normozoosperm	433	971	4,26 (0,1190)	1,88 (0,3900)	2,42 (0,2977)	2,93 (0,2307)

Die Schätzparameter der einzelnen Tests unterscheiden sich aufgrund der spezifischen Gewichtungsfaktoren für die Relationen von für die einzelnen Ausfallzeiten erwarteten und beobachteten Ereignissen.

6.3.2 Ergebnisse der Cox-Modelle

Die Modellierung von Survivalverläufen ohne weitere Kovariablen durch nichtparametrische Verfahren konnte zeigen, dass die Fertilität als einziger Erklärungsfaktor scheinbar nicht geeignet ist, um Sterblichkeitsunterschiede vorherzusagen.

Als semiparametrische Verfahren eignen sich Cox-Modelle immer dann, sobald Einflüsse von erklärenden Variablen mit unterschiedlichen Messniveaus ermittelt werden sollen. Für die Kovariablen müssen dann entsprechende Annahmen formuliert werden, der Baseline-Hazard bleibt dagegen unspezifiziert.

Als Kovariablen werden das Geburtsjahr und der Wohnort herangezogen, welche vermutlich auch einen Einfluss auf die zu erklärende Variable Lebensdauer haben bzw. den potentiellen Effekt von Fertilitätsstatus und Lebensdauer verzerren.

Der Fertilitätsfaktor wird dabei approximativ durch die Variablen Spermienkonzentration sowie durch den Fertilitätsstatus gemäß der WHO-Klassifikation (2010) beschrieben.

Bei der Analyse von kategorialen Variablen wie z.B. dem Fertilitätsstatus können neben den Regressionskoeffizienten auch die relativen Verhältnisse der Hazardneigung subgruppenspezifisch durch die Hazard-Ratios beschrieben werden. Die Auswertung bezieht sich wiederum auf 974 normozoosperme, 255 oligozoosperme und 205 azoosperme Fälle. Bei fünf Fällen konnte keine exakte Zuordnung zum aktuellen Wohnort ermittelt werden.

In den dargestellten statistischen Modellen wird der Regressionskoeffizient mit Coef. beschrieben, SE gibt den Standardfehler an und mit HR wird das Hazard-Ratios bezeichnet. Die Regressionsmodelle folgen stets der Teststatistik des Likelihood-Ratio-Tests (LR) gemäß der χ^2 -Verteilung unter der Berücksichtigung der Freiheitsgrade df.

Tabelle 11: Cox-Modell zum Einfluss der Spermienkonzentration
auf den Hazard mit Geburtsjahr und Wohnort

	Coef	SE	z	P > z	95 % CI	
Spermienkonzentration	-0,00071	0,00068	-1,05	0,295	-0,00204	0,00618
Geburtsjahr	-0,00717	0,00736	-0,97	0,330	-0,02160	0,00726
Wohnort	-0,07081	0,06729	-0,85	0,396	-0,18897	0,07481

n=1399, Log likelihood -2721,5735, LR 2,83 Prob Chi² 0,4179, df=3

Das dargestellte Cox-Modell (Tab.11) mit der Spermienkontration gemessen in Mio./ml als Einflussfaktor auf den Hazard zeigt, dass diese Erklärungsgröße nicht geeignet ist, um mögliche Unterschiede im Hazardverlauf zu erklären.

Für die beiden Kovariablen Geburtsjahr und Wohnort lassen sich ebenfalls keine signifikanten Einflüsse erkennen.

Verstärkt wird die Aussage unter Berücksichtigung des geringen Wertes des Likelihood Ratio Testes (LR 2,83 und Chi² 0,4179). Dabei kann die Nullhypothese von $\beta=0$ nicht verworfen werden, somit ist davon auszugehen, dass die insgesamt eingeschlossenen Kovariablen nicht geeignet sind um einen Effekt für die Sterblichkeit zu erklären.

Tabelle 12: Cox-Modell zum Einfluss des Fertilitätsstatus auf den Hazard
für subfertile und fertile Fälle

	Coef	SE	z	P > z	HR	95 % CI	
subfertil vs. fertil	0,20820	0,10199	2,04	0,041	1,23146*	1,00834	1,50395
Geburtsjahr	-0,00622	0,00736	-0,84	0,130	0,99380	0,97956	1,00824
Wohnort	-0,05339	0,06736	-0,79	0,989	0,94801	0,83076	1,08182

n=1399, Log likelihood -2215,1833, LR 5,78 Prob Chi² 0,1230, df=3

Die Parameterschätzungen der Cox-Regression (Tab.12) zeigen einen (schwach) signifikanten Unterschied ($p=0,041$) in der Hazardneigung zwischen subfertilen und fertilen Fällen. Dabei ist der Hazard der subfertilen Patienten um den Faktor 1,23 erhöht.

Für die beiden eingeschlossenen Kovariablen Geburtsjahr und Wohnort lassen sich ebenfalls keine signifikanten Einflüsse erkennen.

Aufgrund des geringen Wertes des Likelihood-Ratio-Tests (LR 5,78) ist insgesamt davon auszugehen, dass in dem aufgeführten Modell keine Effekte durch die eingeschlossenen Prädiktoren zu erkennen sind.

Tabelle 13: Cox-Modell zum Einfluss des Fertilitätstatus auf den Hazard für azoosperme, oligozoosperme und normozoosperme Fälle

	Coef	SE	z	P > z	HR	95 % CI	
azoosperm vs. normozoosperm	0,17785	0,12940	1,37	0,169	1,19464	0,92703	1,53950
oligozoosperm vs. normozoosperm	0,24198	0,13334	1,81	0,070	1,27377	0,98083	1,65420
Geburtsjahr	-0,00637	0,00737	-0,86	0,387	0,99365	0,97941	1,00810
Wohnort	-0,05264	0,06743	-0,78	0,435	0,94873	0,83127	1,08278

n=1399, Log likelihood -2720,027, LR 5,93 Prob Chi² 0,2047, df=4

Für dieses Cox-Modell (Tab. 13) kann kein statistischer Zusammenhang zwischen dem Fertilitätsstatus und dem Sterberisiko ermittelt werden, was bedeutet, dass im direkten Vergleich sich die Hazards von azoospermen und oligo-zoospermen Fällen gegenüber den normozoospermen Referenzgruppen nicht unterscheiden.

Auch für die beiden anderen Kovariablen Geburtsjahr und Wohnort können keine statistisch signifikanten Effekte ermittelt werden.

Die Konsistenz der Schätzergebnisse für eine Cox-Regression ist nur bei Nichtverletzung der Proportional-Hazard-Annahme gegeben.

Die Überprüfung dieser Annahme erfolgt numerisch durch die Scoretests für die skalierten Schoenfeld-Residuen.

Tabelle 14: Scoretest für die skalierten Schoenfeld-Residuen
für subfertile und fertile Fälle

	rho	chi2	df	Prob>chi2
subfertil vs. fertil	0,05209	1,16	1	0,2816
Geburtsjahr	-0,0281	0,37	1	0,5444
Wohnort	0,03753	0,62	1	0,4323
Global Test		2,1	3	0,5509

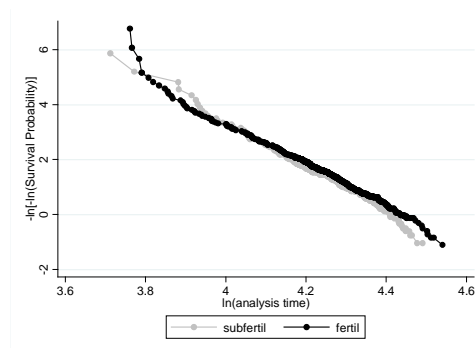
Für den Globaltest (Tab. 14 und Tab. 15) erfolgt die Überprüfung der Proportional-Hazard-Annahme unter der Bedingung der Nullhypothese $H_0: q_j = 0$, eines Steigungskoeffizienten von null für die Schoenfeld-Residuen.

Die Teststatistik (Prob. Chi2 > 0,001) verweist daher auf einer Nichtverletzung der Proportional-Hazard-Annahme.

Tabelle 15: Scoretest für die skalierten Schoenfeld-Residuen
für azoosperme, oligozoosperme und normozoosperme Fälle

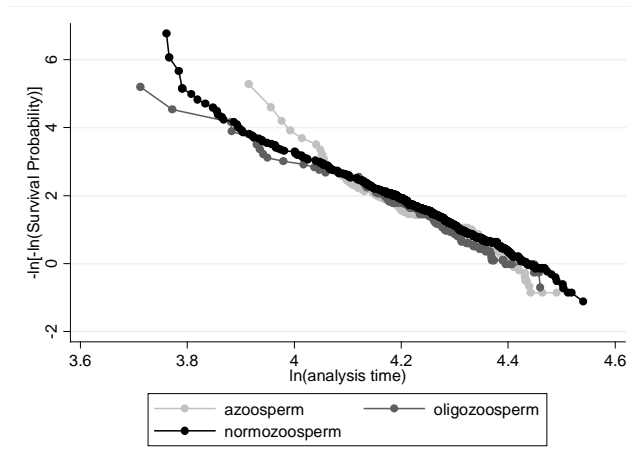
	rho	chi2	df	Prob>chi2
azoosperm vs. normozoosperm	0,05706	1,39	1	0,2391
oligozoosperm vs. normozoosperm	0,04626	0,93	1	0,3357
Geburtsjahr	-0,01413	0,09	1	0,7603
Wohnort	0,0641	1,80	1	0,1793
Global Test		3,75	4	0,4408

Abbildung 13: Überprüfung der Proportional-Hazards-Bedingung
für subfertile und fertile Fälle



Bei Nichtverletzung der Proportionalitätsannahme sollten die erzeugten Kurven in Abb.13 und Abb. 14 möglichst parallel verlaufen und sich in ihrer Form sehr ähneln. Auf Grundlage der dargestellten graphischen Verfahren kann aber nicht eindeutig entschieden werden, ob die Annahme proportionaler Risiken verletzt ist.

Abbildung 14: Überprüfung der Proportional-Hazards-Bedingung
für azoosperme, oligozoosperme und normozoosperme Fälle



Bisher erfolgte eine Modellierung der Cox-Regressionen nur unter Berücksichtigung aller potentiellen Erklärungsfaktoren für den Hazard. Durch den Ausschluss einiger Einflussgrößen ergeben sich „sparsamere“ Modelle, welche oftmals dadurch einen viel größeren Erklärungsgehalt haben und somit als sinnvolle Alternative dem vollständigen Schätzmodell vorzuziehen sind. In diesem Fall wird geprüft ob ein direkter Zusammenhang zwischen dem Survival und der Einflussgröße Fertilität besteht, ohne Berücksichtigung von weiteren potentiellen Einflussgrößen.

6.3.3 Ergebnisse der alternativen Cox-Modelle

Im Vergleich zum vollständigen Modell (siehe Tabelle 16 und Tabelle 17) sind die Schätzergebnisse bei einer Modellierung ohne Berücksichtigung der Variable Wohnort doch sehr ähnlich. Subfertile Fälle haben gegenüber den fertilen Patienten ein erhöhtes Hazard-Risiko (HR: 1,22, $p=0,043$), allerdings ist die Effekt nur schwach signifikant.

Tabelle 16: Alternatives Cox-Modell für subfertile und fertile Fälle

	Coef	SE	z	P > z	HR	95 % CI	
subfertil vs. fertil	0,20600	0,10160	2,03	0,043	1,228757*	1,00690	1,49950

n=1404, Log likelihood -2729,634, LR 4,02 Prob Chi² 0,0449, df=1

Tabelle 17: Alternatives Cox-Modell für azoosperme-, oligozoosperme und normozoosperme Fälle

	Coef	SE	z	P > z	HR	95 % CI	
azoosperm vs. normozoosperm	0,17948	0,12908	1,39	0,164	1,19464	0,92913	1,54106
oligozoosperm vs. normozoosperm	0,23510	0,13247	1,77	0,750	1,27377	0,97576	1,64008

n=1404, Log likelihood -2729,577, LR 4,14 Prob Chi² 0,1265, df=2

Unter Berücksichtigung der Parameterwerte im Cox-Modell für azoosperme, oligozoosperme und normozoosperme Fälle ergeben sich wiederum keine signifikanten Unterschiede in der Hazardneigung. Abermals wird deutlich, das aufgrund des geringen Wertes des Likelihood-Ratio-Tests die Nullhypothese von $\beta=0$ nicht verworfen werden kann, sodass davon auszugehen ist, das in dem Modell keine Effekte durch die verwendeten Kovariablen beschrieben werden.

6.3.4 Parametrische Regressionsmodelle mit Unshared Frailty

Anders als bei den Proportional-Hazards-Modellen von Cox erfolgt eine Spezifikation der Hazardfunktion durch eine entsprechende Verteilungsfunktion. Die Proportional-Hazards-Bedingung bleibt weiterhin aufrechterhalten, weil weiterhin davon ausgegangen wird, dass der Baseline-Hazard multiplikativ von den nicht immer vollständig verfügbaren Kovariablen bestimmt wird. Deshalb sind unbekannte beobachtungsspezifische Effekte möglich, die für die Modellbildung zunächst nicht berücksichtigt werden konnten, aber bei der Untersuchungspopulation heterogene Strukturen erzeugen.

Die Frailty-Eigenschaft wird als unshared modelliert, für den Fall, dass keine gruppenspezifische Korrelationen zu beobachten sind. Die Prüfgröße für die Frailty-Varianz Z wird beschrieben durch den Koeffizienten θ unter Berücksichtigung der Nullhypothese von $H_0: \theta = 0$ und ist gültig für den homogenen Fall.

Die Überprüfung der Nullhypothese $H_0: \theta = 0$ erfolgt dabei mit dem Likelihood-Ratio-Test. In STATA kann der Nutzer bei der parametrischen Spezifikation von Z zwischen der Gamma- und der inversen Gauß-Verteilung entscheiden.

Die Unshared-Fraily-Modelle mit der inversen Gauß-Funktion als Spezifikation konnten nicht berechnet werden, was wohl auf die spezifischen Modelleigenschaften zurückzuführen ist. Die Modellierung von Fraily-Effekten für das Weibull-Modell war nicht möglich, daher wird das jeweilige Basis-Modell dargestellt.

Tabelle 18: Exponential-Modell mit Unshared (Gamma)

Fraily für subfertile und fertile Fälle

	Coef	SE	z	P> z	HR	95 % CI	
subfertil vs. fertil	0,08265	0,10190	0,81	0,417	1,08616	0,88952	1,32627
Geburtsjahr	-0,07002**	0,00534	-13,12	0,000	0,93237	0,92267	0,94218
Wohnort	-0,01500	0,06683	-0,22	0,822	0,985112	0,86417	1,12298
Konstante	130,45360	10,27521	12,7	0,000		85,64335	132,32760
theta	4,18E-08	2,77E-05					
chi2	0,000	1,000					

n=1399, Log likelihood -596,64367 Prob 149,39 Chi² <0,001, df=3,
Signifikanz bei p<0,001

Tabelle 19: Exponential-Modell mit Unshared (Gamma) Fraily für

azoosperme, oligozoosperme und normozoosperme Fälle

	Coef	SE	z	P> z	HR	95 % CI	
azoosperm vs. normzosperm	0,07706	0,12925	0,60	0,551	1,08010	0,83840	1,39150
oligozoosperm vs. normozoosperm	0,08868	0,13286	0,67	0,504	1,09274	0,84222	1,41776
Geburtsjahr	-0,070052**	0,00534	-13,1	0,000	0,93235	0,92263	0,94216
Wohnort	-0,01490	0,06685	-0,22	0,824	0,98521	0,86422	1,12314
Konstante	130,512	10,28570	12,69	0,000		110,352	150,6716
theta	9,33E-08	3,94E-05					
chi2	0,00	1,00					

n=1399, Log likelihood -596,64119, LR 149,39 Prob Chi² <0,001, df=4 Signifikanz bei p<0,001

Tabelle 20: Weibull-Modell für subfertile und fertile Fälle

	Coef	SE	z	P> z	HR	95 % CI	
subfertil vs. fertil	0,19325	0,10150	1,90	0,057	1,21318	0,99433	1,48021
Geburtsjahr	-0,01052	0,00665	-1,58	0,114	0,98954	0,97672	1,00252
Wohnort	-0,05594	0,06713	-0,83	0,405	0,94559	0,82902	1,07856
Konstante	-13,88826	13,67406	-1,02	0,310		-40,68892	12,91241

n=1399, Log likelihood -329,09926, LR 7,23 Prob Chi² 0,0650, df=3,

Tabelle 21: Weibull-Modell für azoosperme, oligozoosperme und normozoosperme Fälle

	Coef	SE	z	P> z	HR	95 % CI	
azoosperm vs. normzoosperm	0,16027	0,12882	1,24	0,213	1,17383	0,92095	1,52562
oligozoosperm vs. normozoosperm	0,23018	0,13294	1,73	0,083	1,25882	0,98375	1,65695
Geburtsjahr	-0,01069	0,00666	-1,61	0,108	0,98937	0,98212	1,00923
Wohnort	-0,05496	0,06722	-0,82	0,414	0,94652	0,82941	1,07922
Konstante	-13,57219	13,68358	-0,99	0,321		-29,34538	23,94612

n=1399, Log likelihood -329,00983, LR 7,23 Prob Chi² 0,1160, df=4

Tabelle 22: Gompertz-Modell mit Unshared (Gamma) Frailty für fertile und subfertile Fälle

	Coef	SE	z	P> z	HR	95 % CI	
subfertil vs. fertil	0,20498	0,10149	2,02	0,043	1,22750*	1,00608	1,49766
Geburtsjahr	-0,00425	0,00694	-0,61	0,541	0,99576	0,98231	1,00940
Wohnort	-0,05638	0,06708	-0,84	0,401	0,94518	0,82874	1,07799
Konstante	-3,04976	13,58417	-0,22	0,822		-29,67423	23,57471
theta	1,52E-06	8,29E-04					
chi2	0,000	1,000					

n=1399, Log likelihood -328,0083, LR 5,30 Prob Chi² 0,1511, df=3

Tabelle 23: Gompertz-Modell mit Unshared (Gamma) Frailty für azoosperme, oligozoosperme und normozoosperme Fälle

	Coef	SE	z	P> z	HR	95 % CI	
azoosperm vs. normzoosperm	0,17003	0,12876	1,32	0,187	1,18534	0,92095	1,52562
oligozoosperm vs. normozoosperm	0,24430	0,13300	1,84	0,066	1,27673	0,98375	1,65695
Geburtsjahr	-0,00443	0,00695	-0,64	0,524	0,99558	0,98212	1,00923
Wohnort	-0,05540	0,06716	-0,82	0,409	0,94611	0,82941	1,07922
Konstante	-2,69963	13,59502	-0,20	0,843		-29,34538	23,94612
theta	2,30E-08	0,0010243					
chi2	0,00	1,00					

n=1399, Log likelihood -328,8732, LR 5,50 Prob Chi² 0,2396, df=4

Bei den aufgeführten vollparametrischen Unshared-Fraily-Modellen konnte nur für die Gompertz(Gamma)-Spezifikation ein signifikanter Unterschied in der Hazardneigung bei fertilen und subfertilen Patienten bestimmt werden. Allerdings gilt dieses Ergebnis unter Vorbehalt, da anhand des geringen Parameterwertes des Likelihood-Ratio-Tests davon ausgegangen werden muss, das möglicherweise keine Effekte im Modell zu beobachten sind, da die Nullhypothese von $H_0: \beta = 0$ nicht abgelehnt werden kann.

Für die Spezifikation mit dem Exponential(Gamma)-Modell und dem Weibull-Modell konnten diese Unterschiede in der Hazardneigung nicht mehr ermittelt werden. Im direkten Vergleich von azoospermen und oligozoospermen Fällen gegenüber der normozoospermen Referenzgruppe ergeben sich keine signifikanten Unterschiede für die Hazards.

Für das Exponential(Gamma)-Modell konnte in beiden Fällen (Tab. 18 und Tab.19) ein statistisch signifikanter Effekt für die Variable Geburtsjahr ermittelt werden.

Die Überprüfung der Nullhypothese von $H_0: \theta = 0$ zeigt, dass diese nicht abgelehnt werden kann, was bedeutet, dass keine unbeobachtete Heterogenität für die Gompertz(Gamma)- und Exponential(Gamma)-Modelle nachweisbar ist.

Eine Entscheidung für die beste Modellanpassung erfolgt nach dem Score-Wert des Akaike-Information-Criterion.

Tabelle 24: AIC für fertile und subfertile Fälle

Verteilung	Log likelihood	k	c	AIC
Exponential (Gamma)	-596,64367	2	1	1199,2873
Weibull	-329,09926	2	2	666,19852
Gompertz (Gamma)	-328,0083	2	2	664,0166

Das Gompertz-Gamma-Modell hat aufgrund der niedrigsten Scorepunkte des AIC die höchste Anpassungstreue.

Tabelle 25: AIC für azoosperme, oligozoosperme und normozoosperme Fälle

Verteilung	Log likelihood	k	c	AIC
Exponential (Gamma)	-596,64119	4	1	1205,2824
Weibull	-329,0083	4	2	672,0166
Gompertz (Gamma)	-328,8732	4	2	671,7464

6.4 Der Einfluss von Lost-to-Follow-up

Neben der Beeinflussung der Schätzergebnisse durch Rechtszensierungen müssen auch jene Fälle behandelt werden, über die keine genauen Angaben über ihren Verbleib bzw. den Zeitpunkt des Ausscheidens aus der Studienpopulation verfügbar sind, sodass der Vitalstatus für diese Fälle sich nicht bestimmen lässt.

Diese so genannten Lost-to-Follow-ups können bei vielen wissenschaftlichen Studien nur bedingt in der Analyse Berücksichtigung finden, da entsprechende Vorinformationen über deren strukturellen Zusammensetzung häufig unvollständig oder nur sehr vage sind.

Dennoch durchlaufen die Lost-to-Follow-ups eine Art Selektionsprozess, der im Fall einer rein zufälligen Systematik vernachlässigt werden kann, da die Reliabilität und Validität der gesamten Studien möglicherweise nicht beeinträchtigt wird.

Wenn für diesen Prozess aber ein konkretes systematisches Muster an Ausfällen zu verzeichnen ist, könnte dieses auch eine mögliche Erklärung für eine feststellbare unbeobachtbare Heterogenität in der Studienpopulation sein.

Durch den Effekt von Lost-to-Follow-ups ist es durchaus möglich, dass bestimmte Personen mit besonders günstigen bzw. ungünstigeren Eigenschaften bezogen auf ihre Hazardneigung bereits vor dem zu beobachtenden Sterbeerignis aus der Studienpopulation ausgeschlossen werden.

Je nach Schweregrad der Beeinträchtigung in der individuellen Reproduktionsfähigkeit und Intensität der Therapieform kann bei einer wiederholten Untersuchung der Spermienparameter das Risiko eines fortzeitigen Ausscheidens vermieden werden, da diese Patienten dann länger unter ärztlicher Beobachtung standen.

Da aber nicht bei allen Patienten mehrere Untersuchungen durchgeführt wurden, konnte jeweils nur die erste Untersuchung für die statistische Auswertung aufgenommen werden, da sonst dieser Zustand die Ergebnisse hätte verzerren können.

Tabelle 26: Einfluss von Lost-to-Follow-up bei subfertilen und fertilen Fällen

Lost-To-Follow-up	Fertilitätsstatus		
	subfertil	fertil	Gesamt
ja	303	585	888
nein	433	975	1408
Total	736	585	2296
Pearson (chi2)	2,8607 (0,091)		

Die Überprüfung mit χ^2 -Statistik (Pearson-Chi² 2,8607, p=0,091 und 3,0684, p=0,216) zeigt, dass die unterschiedlichen Fertilitätszustände nicht durch systematische Ausfälle beeinflusst werden.

Tabelle 27: Einfluss von Lost-to-Follow-up bei azoospermen, oligozoospermen und normozoospermen Fälle

Lost-To-Follow-up	Fertilitätsstatus			
	azoosperm	oligozoosperm	normozoosperm	Gesamt
ja	160	143	585	888
nein	206	227	975	1408
Gesamt	366	370	1560	2296
Pearson (Chi2)	3,0684 (0,216)			

6.5 Wohnortspezifische Unterschiede in der Spermienkonzentration

Genauso kann auch ein Selektionseffekt nach dem Wohnort auftreten, da oftmals davon ausgegangen werden kann, dass die besonders schweren Fälle, die nicht aus dem direkten Einzugsgebiet des Klinikums stammen, eine spezialisierte Fertilitätseinrichtung aufsuchen um sich dort behandeln zu lassen.

Mit dem Mann-Whitney-Test wurde getestet ob, sich die Spermienkonzentrationen von Patienten mit Wohnsitz im Stadtgebiet und Landkreis Marburg-Biedenkopf von den Niveauwerten jener Personen unterscheiden, die ihren Wohnsitz außerhalb des Landkreises haben.

Im direkten Vergleich (Tab. 31) der normozo- und oligozoospermen Fälle ergibt sich für die Marburger eine durchschnittliche Spermienkonzentration von 76,57 Mio./ml, der Wert für die Nichtmarburger ist mit 76,17 Mio./ml etwas geringer.

Tabelle 31: Mann-Whitney-Test für die Verteilung der Spermienkonzentration (normozoosperm und oligozoosperm) nach Wohnort

Wohnort	absolute Häufigkeit	Mean	SD
MR+LK MR	611	76,574	76,3335
außerhalb MR	591	76,167	70,1936
Total	1202	76,3722	73,3257
H_0 : Konzentration (MR) = Konzentration (außerhalb MR)			
Prob> z	0,531		

Der Mann-Whitney-Test ($\text{Prob}>|z|=0,531$) verweist darauf, dass die Nullhypothese identischer Spermienkonzentrationen von Marburgern und Nichtmarburgern nicht abgelehnt werden kann.

Tabelle 32: Mann-Whitney-Test für die Verteilung der Spermienkonzentration (normozoosperm) nach Wohnort

Wohnort	absolute Häufigkeit	Mean	SD
MR+LK MR	493	92,7471	75,9952
außerhalb MR	482	92,0124	68,3783
Total	975	92,3824	72,2782
H_0 : Konzentration (MR) = Konzentration (außerhalb MR)			
Prob> z	0,420		

Auch bei der separaten Betrachtung der normozoospermen Fälle konnten keine signifikanten wohnortspezifischen Unterschiede in der Spermienkonzentration ($\text{Prob}>|z|=0,420$) nachgewiesen werden.

7. Diskussion

In diesem Kapitel soll eine kritische Auseinandersetzung mit den bereits dargestellten Ergebnissen aus dem vorangegangenen Abschnitt erfolgen. Zunächst wird die Qualität des verwendeten Datensatzes hinsichtlich Konsistenz analysiert, wobei mögliche Selektionseffekte und potentielle Verzerrungsgrößen herausgearbeitet werden sollen.

Daran anknüpfend erfolgt eine Diskussion der Studienergebnisse unter Berücksichtigung möglicher Einflüsse durch die spezifischen Modellannahmen sowie eine Aufstellung von möglichen Erklärungsansätzen für unbeobachtete Heterogenität.

7.1 Datenqualität

7.1.1 Selektionseffekt nach Wohnort

Aus den Angaben der Patientenakten lässt sich entnehmen, dass etwa 53% der behandelten Patienten ihren Wohnsitz innerhalb des Stadtgebietes Marburgs oder im Landkreis Marburg-Biedenkopf hatten. Dementsprechend muss aber auch eine Berücksichtigung jener Patienten erfolgen, deren Wohnort sich über das ursprüngliche Einzugsgebiet der Kreisgrenze und der hessischen Landgrenze hinaus erstreckt.

Dabei ist durchaus vorstellbar, dass unter den Fällen außerhalb Marburgs die Häufigkeit an schweren pathologischen Befunden mit geringerer Spermienkonzentration höher ist und diese Patienten im akuten Fall häufiger an spezialisierte Einrichtungen überwiesen werden.

Um eine mögliche Verzerrung identifizieren zu können, wurde bereits vor Beginn der Rechercheaktivitäten in einer Voruntersuchung überprüft, ob sich die Spermienkonzentration von innerhalb Marburgs wohnhaften Patienten von derjenigen jener Personen unterschied, die zum Zeitpunkt der Untersuchung außerhalb des Landkreises Marburgs wohnten.

Die Überprüfung zeigte, dass kein Nachweis von Selektivität besteht, sodass ein Zusammenhang zwischen der Höhe der Spermienkonzentration und dem Wohnort sich nicht bestätigen lässt (vgl. 6.5).

Dabei sollte aber auch erwähnt werden, dass die Wohnorte der Patienten außerhalb Marburgs in eher ländlich geprägten Gebieten ohne eine erkennbare Konzentration von Schwerindustrie liegen, sodass eine Beeinträchtigung durch umweltbedingte Einflussfaktoren doch ausgeschlossen werden kann.

Aufgrund der räumlichen Nähe der einzelnen Herkunftsregionen der untersuchten Personen lässt sich eine vermutete wohnortspezifische Varianz in der Spermienkonzentration ebenfalls nicht begründen.

7.1.2 Selektionseffekt nach Auffindbarkeit

Die Reliabilität und Validität von Untersuchungsergebnissen einer wissenschaftlichen Studie werden häufig durch weitere systematische Selektionseffekte wie z.B. den Grad der Auffindbarkeit während der Follow-up-Phase bestimmt. Dabei ist es durchaus vorstellbar, dass eine bestimmte Gruppe von Fällen mit besonders guten oder schlechten Lebensrisiken systematisch häufiger von Lost-to-Follow-up betroffen ist, sodass deren aktueller Vitalstatus mit einer bekannten Wahrscheinlichkeit häufiger ungeklärt bleibt.

Bis zur Beendigung der Follow-up-Phase konnte für 1408 Personen der Vitalstatus anhand der Informationen aus den Einwohnermeldeämtern und der Krankenkassenkartei bestimmt werden; dieses entspricht einer Responserate von etwa 61% aller erfassten Fälle (2296 Fälle insgesamt). Es galt dabei abzuklären, inwiefern eine bessere Gesundheit, gemessen an der höheren Spermienkonzentration, sich auf die Häufigkeit der Auffindbarkeit auswirke.

Generell lässt sich vermuten, dass Patienten mit schweren pathologischen Befunden eine größere Motivation haben, sich häufiger untersuchen zu lassen, um damit vermeintlich alle Behandlungsmöglichkeiten auszuschöpfen.

Andererseits ist es durchaus möglich, dass bei Patienten mit Azoospermie oder einer niedrigeren Spermienkonzentration der Behandlungserfolg ausbleibt, die ungewollte Kinderlosigkeit weiterhin besteht, die Ehe möglicherweise gelöst wurde, sodass die betroffenen Personen unter diesen Umständen aus dem Untersuchungsgebiet unbekannt verzogen sind.

Daher wurde geprüft, ob sich der Fertilitätsstatus von Patienten mit unbekanntem Vitalstatus von jenen Personen unterscheidet, deren Zustand bekannt war. Die Überprüfung zeigte, dass kein systematischer Unterschied in der Anfälligkeit von Lost-to-Follow-up in Abhängigkeit vom Fertilitätsstatus besteht (vgl. 6.4). Folglich kann in den untersuchten Fällen der Einfluss einer systematischen Selektion weder nach Wohnort noch nach Schweregrad des pathologischen Befundes bestimmt werden. Die untersuchten Fälle werden daher als Zufallsstichprobe innerhalb der Studienpopulation betrachtet.

Allerdings muss auch erwähnt werden, dass der Anteil der in den jüngeren Jahren verstorbenen Fälle potentiell ein wenig unterschätzt sein könnte, da die kooperierende Krankenkasse aus Datenschutzgründen keinen direkten Zugang zu den Archiv-Informationen bereitstellen konnte (Groos, 2005). Für jene Fälle, die bei der Krankenkasse versichert waren und vor der Umstellung auf die elektronische Datenverarbeitung verstorben sind, konnte der Vitalstatus nicht bestimmt werden.

Allerdings konnten mit der Recherche über die Einwohnermeldeämter mehr als 52% aller Fälle direkt identifiziert werden. Der Anteil jener Patienten, die vor 1980 verstorben sind, beträgt etwa 5%, sodass die Unterschätzung mit großer Wahrscheinlichkeit als gering eingestuft werden kann.

Kritisch hervorzuheben sind Risiken von potentiellen Verzerrungen bzw. Verfälschungen, die sich aufgrund von fehlenden Daten nicht eindeutig beantworten lassen. Die vorliegende Untersuchung eines Zusammenhangs von Überlebenszeit und Spermienqualität bezieht bekanntlich nur jene Fälle (insgesamt 1408) ein, für die der Vitalstatus bekannt war. Für die verbliebenen erfassten 888 Fälle konnten während der Follow-up-Phase keine Informationen zur Überlebenszeit ermittelt werden, da diese Fälle aufgrund von Lost-to-Follow-up vor Studienende ausgeschieden sind.

Es ist daher nicht auszuschließen, dass für die gesamte Studienpopulation sich möglicherweise andere statistische Zusammenhänge zwischen den beiden Effektgrößen Fertilitätsstatus und Survival ergeben hätten, wenn für diese Fälle die entsprechende Information zum Vitalstatus vorgelegen hätte. Daher sind auch mögliche Niveaunterschiede in der Exzessmortalität zwischen fertilen und subfertilen Patienten vorstellbar.

Folglich sind eine exakte Interpretation und die Einordnung der Schätzergebnisse nur dann zulässig, wenn dabei nur jene Teilpopulation berücksichtigt wird, bei der die Angaben zum Vitalstatus bekannt sind.

Ebenso ist zu beachten, dass für die Analyse für alle eingeschlossenen Fälle jeweils nur das erste Untersuchungsdatum berücksichtigt wurde. Bei einigen Patienten lagen aber auch Angaben von einer oder mehreren Kontrolluntersuchungen vor.

Somit könnte das Ergebnis der ersten Untersuchung von jenen der Folgeuntersuchungen abweichen, da bei Behandlungserfolg mit einer Verbesserung der Spermienparameter zu rechnen wäre.

Dieses könnte bedeuten, dass bei einige Untersuchungspersonen im Extremfall ein Statuswechsel von fertil zu subfertil und umgekehrt zu beobachten wäre und sich somit die Proportion von fertilen und subfertilen Untersuchungspersonen verändern könnte.

Insgesamt erweist sich ein Vergleich von Spermienparametern aus unterschiedlichen Studienpopulationen als schwierig, weil zum einen die Untersuchungsmethoden der einzelnen reproduktionsmedizinischen und andrologischen Einrichtungen voneinander abweichen können.

Werden Patienten aus unterschiedlichen Behandlungszeiträumen betrachtet, besteht durchaus die Möglichkeit, dass sich die Diagnoseverfahren technologisch weiterentwickelt haben, sodass die Genauigkeit der Befunde bezüglich der Spermiedichte, Morphologie und Motilität innerhalb der Studienpopulation variieren kann.

7.2 Definition von Lebensdauer

Im Abschnitt 5.8 wurde die Lebensdauer definiert als gesamte Lebensspanne ab dem jeweiligen Geburtsdatum bis zum Sterbedatum bzw. Zensierungstag (31.12.2006).

Eine Betrachtung der gesamten Lebensdauer vom Geburtsdatum bis zum Eintreten des Todes erweist sich für die Fragestellung als schwierig und problematisch, da im Datensatz die betreffenden Angaben unvollständig sind, da die Zeitspanne bis zum Zeitpunkt des ersten Untersuchungstermins nicht direkt beobachtet werden kann.

Es muss daher eine ausreichende Berücksichtigung von Linkstrunkierungen erfolgen, da einige Patienten bereits einige Zeit überlebt hatten, bevor sie in die Studie aufgenommen wurden.

Zudem können die Eintrittszeitpunkte unter Beobachtung und demzufolge auch die Länge der Follow-up-Phase innerhalb der Untersuchungspopulation variieren. Eine entsprechende Berücksichtigung dieser Variation erfolgt unter Modellierung des delayed entry, wie in Abschnitt 5.8 beschrieben.

Ebenso ist eine entsprechende parametrische Modellschätzung mit der sehr verbreiteten Gompertz-Spezifikation (Gompertz, 1825) für die gesamten Lebensspanne problematisch, weil mit ihr nur jener Teil der Sterblichkeit im Erwachsenenalter (exponentiell steigende Hazards) erklärt werden kann; der Teil mit einer exponentiell sinkenden Hazardfunktion im Kindesalter bleibt dabei unberücksichtigt.

Dieses Problem greifen ebenfalls Engelman et al. 2009 in einer neuen wissenschaftlichen Arbeit auf und entwickelten das so genannte Siler-(Gamma)-Modell für Overall-Survivalverläufe.

Das Siler-Modell, das ursprünglich als Competing-Risk-Modell bei der Untersuchung von evolutionär bedingten Anpassungsstrategien von Tieren an die gegebenen Umweltbedingung Anwendung findet (Siler, 1979), erschließt eine Dekomposition des Overall-Hazards in drei Teilkomponenten:

1. Eine Phase mit sinkendem Hazard bei positiver Adjustierung des Individuums beschreibt jenen Abschnitt im Kindesalter.
2. Eine Phase mit konstantem Hazard ohne Adjustierung beschreibt jenen Abschnitt im jungen Erwachsenenalter.
3. Eine Phase mit steigendem Hazard bei negativen Anpassungsmechanismen durch das Individuum im mittleren bis hohen Erwachsenenalter gekennzeichnet ist.

Die Innovation dieses Ansatzes liegt in der Formulierung von altersspezifischen Frailty-Variablen, um damit altersbedingte Zufallseffekte identifizieren zu können.

Das Siler-(Gamma)-Modell lässt sich folglich definieren durch:

$$\lambda(t) = \frac{\alpha_1 e^{-\beta_1 X} \alpha_2 e^{\beta_2 X} + \alpha_3}{1 + \sigma^2 \left(\frac{-\alpha_1}{\beta_1} (e^{-\beta_1 X} - 1) + \frac{\alpha_2}{\beta_2} (e^{-\beta_2 X} - 1) + \alpha_3 X \right)}$$

Die Gültigkeit des Siler-(Gamma)-Modells konnte bereits für die schwedischen Frauenkohorten von 1845–1915 mit den entsprechenden altersspezifischen Frailties nachgewiesen werden (Engelman et al., 2009).

Ob einer Unter- bzw. Überschätzung des Populations-Hazards für den gesamten Lebensverlauf durch die Gompertz-Funktion erfolgt, kann nur im direkten Vergleich mit einer entsprechenden Modellierung durch das Siler-(Gamma)-Modell nachgewiesen werden; dazu müssten aber die entsprechenden Angaben zum Survivalverlauf für die Phase vor dem eigentlichen Beobachtungsbeginn (Zeitpunkt bei Beobachtungsbeginn) bekannt sein.

7.3. Die Auswirkungen der betrachteten Hypothesen auf die Wahl des Studiendesigns

Der Fokus der vorliegenden Untersuchung liegt auf der Überprüfung des Zusammenhangs zwischen dem Fertilitätsstatus und der Lebensdauer von Männern. Es wird zunächst unterstellt, dass die Fertilität als ein möglicher Erklärungsfaktor für potentielle Sterblichkeitsunterschiede gesehen werden kann. Dabei ist durchaus vorstellbar, dass ein Selektionsprozess wirkt, bei dem jene Individuen herausselektiert werden, die dabei besonders durch schlechte Lebensrisiken gekennzeichnet sind.

Der Fertilitätsstatus kann, wie in Abschnitt 7.8 beschrieben, durch äußere Umwelteinflüsse, aber auch durch individuelle Verhaltensweise beeinträchtigt werden.

Die Fertilität kann daher auch als ein Indikator für den allgemeinen Gesundheitszustand gesehen werden, weil angenommen wird:

1. dass sich nur jene Individuen reproduzieren können, die auch über ein gewisses Maß an körperlicher Fitness verfügen.
2. Erkrankungen der Geschlechtsorgane (u.a. Erektionsstörungen) sowie gesundheitliche Beeinträchtigungen des Verdauungstraktes (u.a. Darminfektionen), die die Reproduktionsfähigkeit und damit die Lebensqualität beeinträchtigen können.

Der Fertilitätsstatus wird dabei approximativ durch die Spermienkonzentration beschrieben. Diese ist wiederum gekennzeichnet durch einen gemessenen Laborwert der andrologischen Zentren und kann nur ein mögliches Potential an Fertilitätseinschränkungen abschätzen, es aber nicht vollständig messen. Die Entscheidung, ob eine Fertilitätsstörung entsprechend vorliegt oder nicht, erfolgt häufig über Einschätzung durch die Referenzwerte der WHO (2010). Diese weist in ihren Ausführungen darauf hin, dass eine reine Beurteilung nach Spermioogramm-Befund nicht ausreicht, um eine mögliche Infertilität exakt zu diagnostizieren bzw. auszuschließen.

Psychische und andere physische Komponenten, die den hormonellen Regelkreis betreffen, können allgemein nicht durch die vorliegenden Spermioogramm-Informationen geklärt werden.

Die Bestimmung des Fertilitätsstatus nur an diesem einen Merkmal ist daher unzureichend. Die Spermienkonzentration ist als exakter Prädiktor für Fertilität geeignet, sobald ein Zusammenhang mit der Schwangerschaftsrate (van Weert et al. 2004, 2008) bei der Partnerin oder der so genannten Baby-Take-Home-Rate, der Erfolgsquote verschiedener Behandlungsmethoden zur Erfüllung des Kindeswunsches, besteht.

Für die Modellierung des Zusammenhangs von Fertilität und dem Survival bei Männern wird dabei auf vollparametrische Frailty-Modelle zurückgegriffen, weil dadurch mögliche heterogene Einflüsse berücksichtigt werden sollen, welche die individuelle Gesundheit der Patienten betreffen, aber anhand der fehlenden Informationen über potentielle Vorerkrankungen nicht direkt gemessen werden können.

Die reine statistische Überprüfung eines solchen Effektes mit Hilfe einer Zufallsvariable, um damit diese möglichen Einflüsse unbeobachteter Heterogenität zu erklären, führt natürlich auch einige Probleme mit sich.

Bei einem positiven Nachweis insbesondere bei Shared-Frailty-Modellen (Yashin et al., 1995a) kann möglicherweise nicht ausgeschlossen werden, dass diese Diagnostik möglicherweise zu einem falschen Ergebnis oder zu einer irreführenden Interpretation führt. Es handelt sich dabei stets um eine unbeobachtete Einflussgröße, die häufig unabhängig von der vorliegenden Datengrundlage geschätzt wird.

Eine ausführliche Erläuterung der Modelleinschränkungen von parametrischen Frailty-Modellen erfolgt in Abschnitt 7.6.

Ebenso ist vorstellbar, dass sowohl der Fertilitätsstatus als auch die Sterblichkeit von einer gemeinsamen unbeobachteten Einflussgröße (z.B. durch genetische Einflüsse) bestimmt werden.

Eine entsprechende Modellierung könnte dabei beispielsweise mit Hilfe eines Shared-Frailty-Modells erfolgen. Patienten mit ähnlichen Spermogramm-Befunden würden sich dann gemäß der Modellannahme die gleichen Lebensrisiken „teilen“ und daher voraussichtlich auch ähnliche Sterberisiken aufweisen. Vorausgesetzt, es existiert so ein gemeinsamer Effekt, dann müsste dieser auch mit den vorliegenden Daten auch eindeutig identifiziert werden können.

Bei den Shared-Frailty-Modellen ergibt sich das Problem, dass sich diese nur auf einziges gemeinsames Attribut beziehen und daher auf Unterschiede innerhalb einer Untersuchungsgruppe nur unzureichend eingehen.

Daher empfiehlt sich die Modellierung mit Correlated-Frailty-Ansätzen (siehe u.a. Wienke et al. 2003; Hens et al. 2009).

Diese sind in ihrer Modellkonzeption wesentlich komplexer, können aber sowohl positive als auch negative Effekte bezüglich gemeinsamer individueller Merkmale bei den Probanden beschreiben und daher auch Unterschiede in der Frailty-Neigung berücksichtigen.

Eine solche Modellierung ist für diese Arbeit nicht vorgesehen, könnte aber, wenn die entsprechenden individuellen Informationen beispielsweise zu den Vorerkrankungen oder der genetischen Exposition vollständig vorlägen, entsprechend durchgeführt werden. Sind die genetischen Informationen der Probanden vorhanden und aussagekräftig, dann wäre es eventuell möglich zu bestimmen, ob eine Fertilitätsstörung des Vaters an die männlichen Nachkommen weitergegeben würde.

7.4 Bewertung der Regressionsschätzungen mit dem Faktor Gesamtfertilität

Für die Beurteilung der Gesamtfertilität wurde in der Arbeit von Seydel (2010) mit Hilfe einer Hauptkomponentenuntersuchung ein Index gebildet, welcher sich aus den Teilkomponenten Spermienkonzentration, Spermiengesamtzahl, Motilität und Morphologie zusammensetzt. Aus der Dokumentation wird nicht eindeutig ersichtlich, ob es sich dabei um einen additiven oder multiplikativen Index handelt. Ebenso bleibt unklar, ob bzw. wie die einzelnen Teilkomponenten gewichtet werden.

Zwar wurde der Einfluss dieser Teilkomponenten auf die Gesamtfertilität bestimmt, die potentiellen Effekte der einzelnen Faktoren auf die Sterblichkeit wurden nicht überprüft.

So ist zu vermuten, dass erst durch die Indexbildung und die damit verbundene Kumulierung der kleineren Teileffekte sich der Gesamteffekt für die Sterblichkeit verstärkt hat.

Die Regressionsanalysen mit dem neu gebildeten Gesamtindex konnten auch nur für jene Teile der Stichprobe durchgeführt werden, bei denen die Angaben für alle relevanten Faktoren der Fertilität vollständig sind.

7.5 Diskussion der Ergebnisse

Die vorliegende wissenschaftliche Arbeit hatte, wie bereits erwähnt, als Zielsetzung eine Untersuchung der Überlebenszeitverläufe von subfertilen und fertilen Männern unter der Berücksichtigung von Zufallseffekten. Es erfolgte eine Analyse mit nicht-parametrischer, semiparametrischer und parametrischer Modellierung, im Hinblick auf Unterschiede im Survivalverlauf von fertilen und subfertilen Männern bzw. zwischen normozoospermen, oligozoospermen und azoospermen Männern.

Ein wesentlicher Bestandteil dieser Arbeit ist auch die Überprüfung der semiparametrischen und parametrischen Modellrechnungen auf unbeobachtete Heterogenität, welche anhand einer Zufallsvariablen der Frailty identifiziert werden soll.

Neben der deskriptiven Darstellung und der kritischen Einschätzung der einzelnen Studienergebnisse sollen auch aber auch methodische Probleme mit dem Fokus auf die Identifikation der Frailty-Varianz sowie potentielle Erklärungsansätze für die unbeobachtete Heterogenität herausgearbeitet werden.

7.5.1 Ergebnisse der nichtparametrischen Verfahren

Die graphische Überprüfung der Survivalverläufe der subfertilen und fertilen Patienten erfolgte mit Überlebenskurven des Kaplan-Meier-Schätzers und der kumulierten Hazardfunktion von Nelson-Aalen.

Die Ergebnisse erwiesen sich als nahezu identisch und zeigten nur marginale Unterschiede im Survival- bzw. Hazardverlauf. Die Sterbefunktionen von subfertilen und fertilen Fällen sind im Verlauf daher kaum voneinander zu unterscheiden.

Auch bei dem Vergleich der Survivalverläufe von normozoospermen, oligozoospermen und azoospermen Patienten zeigten sich nur geringfügige Unterschiede.

Eine rein optische Beurteilung von Survivalverläufen mit nichtparametrischen Verfahren ohne Berücksichtigung von dynamischen Konfidenzbändern, z.B. Equal-Precision-Band oder Hall-Wellner-Band, die sich im Laufe der Beobachtungszeit an die sinkende Zahl an Ereigniszeiten anpassen, ist daher nicht ausreichend.

Im Extremfall kann dieses zu systematischen Fehleinschätzungen und Schlussfolgerungen führen (Hall und Wellner, 1980; Nair, 1984; Lin, 1997). Deshalb wurden auch die Teststatistiken Log-rank, Wilcoxon, Tarone-Ware und Peto-Peto-Prentice herangezogen.

Keines der beschriebenen Testverfahren konnte einen Unterschied im Overall-Survival-Verlauf ermitteln, sodass zunächst davon auszugehen ist, dass der Fertilitätsstatus keine besonders geeignete Variable ist, um Unterschiede im Survivalverlauf zu erklären. Die dargestellten nichtparametrischen Verfahren können nur eine weitere Erklärungsgröße, z.B. den Fertilitätsstatus, mit berücksichtigen.

Mögliche Effekte von Interaktionen, wie mit dem Untersuchungsjahr oder Wohnort, können zunächst nicht beachtet werden.

7.5.2 Semiparametrische Schätzverfahren (Cox-Regression)

Bei den Cox-Modellen ließen sich signifikante Unterschiede in der Hazardneigung für die fertilen und subfertilen Patienten ermitteln, diese sind aber insgesamt nicht sehr groß. Die Hazards-Ratios der azoospermen und oligozoospermen Fälle, bezogen auf die Normozoospermen als Referenz, zeigten keine signifikanten Unterschiede.

Die Konsistenz der Schätzergebnisse für die Cox-Modelle wird beeinflusst durch die Gültigkeit der Proportional-Hazards-Annahme, die sich sowohl graphisch als auch rechnerisch überprüfen lässt.

Die rechnerische Überprüfung mit der Teststatistik der skalierten Schoenfeld-Residuen als eine Funktion über die Zeit (Schoenfeld, 1982) konnte keinen Nachweis für eine Verletzung der Proportional-Hazards-Annahme erbringen. Das Ergebnis der graphischen Überprüfung ist im Vergleich dazu weniger eindeutig.

Dies hängt möglicherweise damit zusammen, dass die Unterschiede in den Hazards-funktionen zu gering sind, sodass sich die Hazards-Verläufe nicht derartig voneinander unterscheiden lassen. Insbesondere bei kategorialen Auswertungen mit mehr als drei Ausprägungen und marginalen Unterschieden in der Sterbeneigung können die einzelnen Verläufe der Teilmerkmale nicht immer eindeutig voneinander unterschieden werden.

Die Modellschätzungen der Cox-Modelle sind aus zweierlei Hinsicht kritisch zu bewerten:

1. Besteht immer das Problem der Gültigkeit der Proportional-Hazards-Annahme?
2. Erfolgt die Schätzung für unbekannte Regressionskoeffizienten mit Hilfe des partiellen Likelihoods?

Eine detaillierte Behandlung des Identifikationsproblem und der Gültigkeit der Proportional-Hazards-Annahme erfolgt im Abschnitt 7.6.1.

7.5.2 Ergebnisse der parametrischen Regressionsmodelle mit Frailty-Varianz

Die Ergebnisse der vollparametrischen Verfahren zeigten einen (schwach) signifikanten Unterschied in der Sterblichkeit für subfertile und fertile Patienten unter Berücksichtigung der Gompertz(Gamma)- Spezifikation. Allerdings gilt dieses Ergebnis unter Vorbehalt.

Wird die Teststatistik des Likelihood-Ratio-Test herangezogen, so kann die Nullhypothese H_0 von $\beta=0$ nicht verworfen werden. Es ist somit davon auszugehen, dass mit großer Wahrscheinlichkeit keine statistischen Effekte durch die im Modell berücksichtigten Kovariablen zu ermitteln sind.

Die Hazardfunktionen von Azoospermen und Oligozoospermen unterscheiden sich unabhängig von der Parametrisierung nicht von dem Niveauwert der normozoospermen Vergleichsgruppe. Bis auf die Modelle mit der Exponential(Gamma)-Spezifikation konnte ein statistischer Zusammenhang der Kovariable Geburtsjahr bestimmt werden.

Dieses Ergebnis ist wohl eher auf die spezifischen Modellannahmen aufgrund der Parametrisierung durch die Exponential-Funktion zurückzuführen, da dieser Effekt weder mit der Gompertz(Gamma)- noch mit der Weibull-Spezifikation ermittelt werden konnten.

Die statistische Überprüfung auf die Existenz von unbeobachteter Heterogenität ist bei allen Unshared-Fraily-Modellen ohne Behandlung einer gruppenspezifischen Fraily-Varianz immer negativ. Die Schätzungen für Unshared-Fraily-Modelle bei einer Fraily mit einer inversen Gauß-Funktion sowie das Weibull-Modell mit Gamma-Spezifikation für die Fraily-Varianz konnten nicht berechnet werden; dies ist möglicherweise ein Hinweis auf das Spezifikationsproblem bei parametrischen Modellen.

Bedingt durch die spezifischen Modellannahmen und die Datenstruktur lassen sich bestimmte Modellverfahren nicht umsetzen. Unter Berücksichtigung des AIC konnte für die Gompertz(Gamma)-Spezifikation insgesamt der geringste Parameterwert bestimmt werden, sodass diese innerhalb der Betrachtung der vollparametrischen Frailty-Modelle die beste Modellanpassung liefert.

Die Gompertz(Gamma)-Spezifikation ist bei der Modellierung von Überlebenszeiten ein weitverbreitetes Verfahren. Dabei wird unter Berücksichtigung eines multiplikativen Frailty-Terms die gültige Beziehung aus der Varianz der Lebensspanne und dem Gompertz-Slope-Parameter (log Mortalität) gelöst (Trivon, 2010; Tuljapurkar und Edwards, 2011).

7.5 Methodische Probleme der Frailty-Ansätze

Wie jeder Modell-Ansatz sind auch Frailty-Modelle von gewissen methodischen Einschränkungen betroffen. Finden diese Schwächen keine ausreichende Beachtung, entwickeln sich eine Vielzahl von unbegründeten Behauptungen sowie fehlerhafte Anwendungen und Umsetzungen dieser Modelle, welche dann zwangsläufig zu eher zweifelhaften Resultaten und Interpretationen führen.

Auf eine fehlende kritische Auseinandersetzung mit den Modellrestriktionen des Frailty-Konzepts verweisen Yashin et al. (2001) in einem Aufsatz zu Hidden Frailty: Myths and Reality.

7.6.1 Identifikationsproblem der Proportional-Hazards-Annahme

Die Hazard-Funktion erlaubt eine umfassende Darstellung der Lebenszeiten der Individuen und der Einflussgrößen (z.B. Krankheit, Alter), welche diese maßgeblich mitbestimmen. Sie lässt sich im Wesentlichen in zwei Komponenten aufgliedern:

1. Die Baseline-Hazard-Funktion ist eine Effektgröße, die sich zwar über die Zeit verändert, aber nicht durch die Kovariablen beeinflusst wird. Optional kann eine Spezifikation durch eine statistische Verteilungsfunktion erfolgen, was aber bei bestimmten Modell-Annahmen (z.B. Proportional-Hazards-Modell nach Cox) nicht immer erforderlich ist.
2. Die Funktion der Kovariablen beschreibt, wie diese den Hazard beeinflussen. Für das Proportional-Hazards-Modell nach Cox (1972) sind gemäß den Modell-Annahmen die Hazard-Ratios über die Zeit konstant und daher von der Überlebenszeit unabhängig.

Im Cox-Modell wird, wie bereits erwähnt, die Hazardfunktion nur durch die Regressionskoeffizienten der Kovariablen bestimmt, sodass die Spezifikation des vollständigen Modells nur von der Verteilungsfunktion der Kovariablen abhängt. In diesem Fall wären die Verteilungsfunktionen der Überlebenszeit und der Kovariablen identisch.

Die Kovariablen können also signifikant die Parameter für die Funktion der Überlebenszeit beeinflussen.

Dabei ist darauf zu achten, dass bei der Vielzahl von Einflussgrößen einige nur unvollständig gemessen werden bzw. sogar bewusst oder unbewusst nicht mit in die Analyse aufgenommen werden.

Eine derartige Vorgehensweise kann diverse Gründe haben:

1. Es existieren einfach zu viele Kovariablen, sodass es bei der Modellbildung nahezu unmöglich ist, alle relevanten Faktoren zu berücksichtigen.
2. Aus eher praktischen Gründen werden einige wichtige Kovariablen bewusst nicht beachtet, weil diese nur schwer erfasst werden können und somit möglicherweise noch ausstehende Arbeitsabläufe bei der Datengenerierung dadurch verlangsamt würden.
3. Bestimmte Einflussgrößen lassen sich nicht direkt als potentielle Kovariablen beschreiben.

Genetische Faktoren, Krankheiten mit relativ langen Latenzzeiten und andere unbekannte Faktoren lassen sich somit als Kovariablen kaum erfassen. Bei diesen Effektgrößen handelt es um unbeobachtete Kovariablen.

Die Validität eines Regressionsmodells lässt sich immer dann anzweifeln, sobald unbekannte Komponenten existieren, aber nicht richtig als Erklärungsfaktoren beobachtet werden können. Darüber hinaus unterscheiden sich Individuen bezüglich ihrer individuellen Anfälligkeit gegenüber bestimmten Krankheiten oder dem Tod.

In diesem Fall wird in der Fachliteratur sehr häufig der Begriff Frailty zitiert (u.a. Abbring und van den Berg, 2006; Butt und Haber, 2004; Chuang et al., 2005; Duchateau und Jansen, 2008; Finkelstein und Esaulova, 2006; Vaupel, 1979; Wienke, 2006, sowie Yashin et al., 2008), ein Zustand von nicht beobachtbarer Heterogenität oder Zufallseffekte bei den beobachteten Individuen, welche dann aber eine Veränderung der Survival-Eigenschaften für die gesamte Population zur Folge haben.

Zur Vereinfachung der Modellannahmen könnte man sich dazu entschließen, diese unbekannten Größen für die Analyse zu vernachlässigen, indem sie nur als Fehlerterm bei der Regressionsanalyse mit modelliert würden. Eine separate Kontrolle dieser Einflüsse ist dadurch nicht mehr erforderlich.

Diese Maßnahme würde zwar die Berechnung wesentlich vereinfachen, dennoch ist bei einer möglichen Heterogenität von nichterfassbaren Kovariablen mit einer Veränderung der Variabilität für die gesamte Population zu rechnen.

Welche Auswirkungen unbeobachtete Heterogenität auf die Parameterschätzung der Survival-Funktion hat, wurde sehr umfangreich von Henderson und Oman (1999) untersucht. So konnte erneut gezeigt werden, dass die traditionelle Homogenitätsannahme nur in modifizierter Form angewendet werden kann und daher eher heterogenere Strukturen auf der Individualebene mit jeweils unterschiedlichen Basis-Risiken, aber für die identischen Populationen, zu beobachten (Keyfitz, 1978; Vaupel et al., 1979) sind.

Es existiert demzufolge eine Varianz bezüglich der Anfälligkeit (Frailty) gegenüber bestimmten Ereignissen wie dem Erkrankungsrisiko bzw. dem Risiko zu versterben. Die Individuen unterscheiden sich also bezüglich ihrer Frailty-Eigenschaften.

Jene Individuen mit größeren Anfälligkeiten gegenüber dem Ereignis werden mit einer bekannten Wahrscheinlichkeit eher aus der Risikopopulation ausscheiden oder versterben als jene Personen, die schon zu Beobachtungsbeginn ein deutlich geringeres Frailty-Risiko haben.

Sobald die Zusammensetzung der Untersuchungspopulation zu heterogen ist, dieses aber nicht mit den vorhandenen Kovariablen direkt kontrolliert werden kann, könnten möglicherweise auch alle Schätzungen für die Hazardfunktion verzerrt sein.

Dieses Problem lässt sich lösen, indem Frailty-Modelle zur Evaluierung von unbeobachteter Heterogenität herangezogen werden.

Erstmals wurden solche Modelle für bivariate Fragestellungen von Clayton (1978) etabliert. Der Frailty-Begriff wurde aber erst explizit in einer Publikation von Vaupel et al. (1979) im Zusammenhang mit demographischen Sterbetafeln benannt. Innovativ ist dabei die Modellierung eines nicht messbaren multiplikativen Skalen-Faktors, der Frailty, und dessen Einfluss auf den Baseline-Hazard.

Folglich kann die Hazard-Funktion in drei multiplikative Komponenten untergliedert werden:

1. einen zufälligen Frailty-Effekt (Skalenfaktor umfasst alle separaten Einflüsse von unbeobachteten Kovariablen auf die Hazard-Funktion)
2. Baseline-Hazard,
3. Funktion der Kovariablen.

Die Werte des individuellen Hazards bestimmen die Frailty-Eigenschaften innerhalb der Untersuchungspopulation mit $\lambda = 1$, alle Individuen unterscheiden sich nicht bezüglich der Frailty-Neigung, $\lambda > 1$; einige Individuen sind anfälliger gegenüber einem bestimmten Ereignis und scheiden dadurch früher aus der Risikopopulation aus als andere, $\lambda < 1$ einige Individuen bleiben länger einem bestimmten Risiko ausgesetzt als andere, scheiden aber nicht frühzeitig aus der Population aus.

Weitere Modellannahmen sind:

1. Überlebenszeit ist unabhängig von der gegebenen Frailty-Annahme;
2. Hazard für jedes Survivalereignis entspricht dem Proportional-Hazard-Modell;
3. Frailty-Variable und Kovariablenfunktion wirken multiplikativ auf den Baseline-Hazard;
4. Frailty-Term und Kovariablen sind voneinander unabhängig;
5. Frailty-Koeffizienten sind innerhalb der Untersuchungsgruppe identisch.

Der Frailty-Term bleibt zunächst weiter unbestimmt, für die Schätzung der Survival- und Dichtefunktion kann eine Spezifikation durch eine parametrische Verteilungsfunktion erfolgen.

Wenn aber die Verteilungsannahmen für diesen unbeobachteten Skalenfaktor nicht erfüllt sind, sind die Schätzungen für das gesamte Frailty-Modell sehr unpräzise und müssten daher angezweifelt werden.

Nur eine genaue Kenntnis der Verteilungsfunktion für die Frailties führt letztendlich zu exakten Ergebnissen, wobei dies aus rein praktischer Sicht nicht immer erfüllbar ist.

7.6.2 Das Spezifikationsproblem parametrischer Frailty-Modelle

Zur Identifikation eines Modells sind oftmals Annahmen über die statistische Verteilungsfunktion notwendig. Frailty-Modelle können jegliche Formen einer parametrischen Verteilungsfunktion annehmen. Dabei scheint es häufig aus heutiger Sicht eine gemeinsame „Absprache“ vieler Autoren zu geben, dass eine Erklärung für die Auswahl einer spezifischen Funktion für das Modell nicht notwendig sei.

Bei demografischen Fragestellungen werden häufig Gompertz-Makeham, in epidemiologischen bzw. ökonometrischen Modellierungen eher die Weibull-Funktion für die Spezifikation des Baseline-Hazards angewendet.

Die Spezifikation der Verteilungsfunktion für die Frailties erfolgt ebenfalls eher nach praktischen Kriterien. Leicht zu modellierende Annahmen werden dabei eher bevorzugt.

Die Gamma-Spezifikation ist eine der am meisten verwendeten Verteilungsfunktionen, die für die Modellierung von Frailty-Eigenschaften (Vaupel et al., 1979; Oakes, 1982; Yashin and Iachine, 1995; Hougaard, 2000; Wienke et al., 2002, 2003) genutzt werden kann. Warum ausgerechnet diese Verteilungsfunktion und keine andere parametrische Form anzuwenden ist, lässt sich weder mit biologischen noch mit empirischen Argumenten begründen.

Vereinfacht kann das Gamma-Modell unmittelbar durch eine Laplace-Transformation hergeleitet werden. Erst später feststellbare Abhängigkeiten innerhalb des Modells lassen sich durch die Annahme der Gamma-Frailties relativ frühzeitig erkennen.

Neben diesen benannten Vorteilen ergeben sich aber auch methodische Schwierigkeiten.

Die ursprüngliche zweiparametrische Schätzung für die Modellidentifikation muss einparametrisch erfolgen, da für die Modellvoraussetzung der Skalenparameter und der Formenparameter identisch sein müssen.

7.6.3 Die Frailty-Varianz als alleiniger Indikator für unbeobachtete Heterogenität

Sehr häufig wird behauptet, dass die Varianz in Frailty-Verteilung alleinig die Heterogenitätseigenschaften in der Untersuchungspopulation widerspiegele. Die Frailty ist aber überwiegend eine nichtbeobachtbare Effektgröße und wird erst durch die Spezifikation durch eine Verteilungsfunktion messbar, sodass die eigentlichen Schätzparameter des gesamten Modells nur von der parametrischen Form der unterstellten Verteilungsfunktion abhängen.

Yashin et al. (2001) konnten anhand von Survivaldaten einer weiblichen schwedischen Studienkohorte des Jahres 1862 beweisen, dass unterschiedliche Verteilungsannahmen über Hazard und Frailty bei einem identischen Datensatz zu sehr unterschiedlichen Ergebnissen führen können.

Der Einfluss bzw. das Ausmaß von unbeobachteter Heterogenität kann mutmaßlich von dem Identifikationsproblem des Regressionsmodells bestimmt werden.

7.6.4 Die Modellvoraussetzungen von Shared-Frailty-Ansätzen

Mitunter finden Shared-Frailty-Modelle (ua. Liu et al. 2004, Hens et al. 2009, Jonker et al., 2009, Petersen et al. 2010) in der Praxis u.a. bei Linkage- und Survivalanalyse eine breite Verwendung.

Shared-Frailty-Modelle folgen dabei der Logik eines gemeinsamen Risikos, bei dem Individuen des gleichen Clusters die gleiche Chance entwickeln von einem unbekannten Risikofaktor betroffen zu sein. Nahezu jede Untersuchungspopulation lässt sich in einzelne oder mehrere Cluster aufteilen, bei denen für das jeweilige Cluster die identischen Frailty-Eigenschaften gültig sind. Die Mitglieder eines jeden Clusters besitzen daher die gleichen Frailty-Parameter, deshalb werden diese auch als Shared Frailty bezeichnet.

Für den multivariaten Fall, bei bekannten Shared Frailties, ist die beobachtete Lebenszeit unabhängig und wird nur durch die Clusterzugehörigkeit bestimmt.

Die Schwierigkeit liegt dabei in der Identifikation von Clustern mit Individuen mit besonders variablen Eigenschaften, die somit keine gemeinsamen Lebensrisiken besitzen.

Insbesondere bei Shared-Frailty-Modellen kann oft eine „Überidentifikation“ entstehen, sobald für diese Modelle zu viele Annahmen (Proportional-Hazards-Bedingung für die Kovariablen, konditionale Abhängigkeit der Schätzparameter, endlicher Mean für die Frailty-Verteilung) formuliert werden, sodass sich häufig diese Modelle nicht mehr auf reale Datensätze anwenden lassen.

Die Varianz der Shared Frailty wird häufig als Messinstrument für die Abhängigkeit zwischen zwei Ereigniszeiten genutzt. Wenn die Frailty-Varianz einen Wert von null annimmt, dann sind die beiden Ereigniszeiten voneinander unabhängig und werden nur durch die Kovariablen bestimmt. Durch die bedingte Varianz (bestimmt durch die parametrische Verteilungsfunktion) entsteht eine gewisse Abhängigkeit zwischen diesen Vergleichsgrößen, wobei es durchaus zweifelhaft ist, Abhängigkeitsverhältnisse zwischen Individuen zu bestimmen, die in ihrer Struktur doch sehr stark variieren.

Bei den Shared-Fraily-Modellen werden die Frailty-Terme als relative Risiken beschrieben, die innerhalb eines Clusters identisch sind. Das bedeutet auch, dass die Frailty-Eigenschaften sowohl für das gesamte Cluster als auch für die einzelnen Individuen zutreffen.

Dabei sollte aber auch klar sein, dass eine Shared Frailty im bivariaten Kontext nur jenen Teil des Modells beschreibt, bei dem sich die Individuen bestimmte gemeinsame Merkmale „teilen“; die ungleichen Eigenschaften werden dann mit der jeweils unterstellten Hazard-Funktion modelliert (Yashin et al., 1995a).

Die Hazard-Funktion ist diesem Fall für alle Individuen gleich, sodass diese Modelle letztendlich nur einen einzigen Schätzparameter benötigen. Dieses ist aber nicht mehr möglich, sobald die betrachteten Individuen in der Realität doch zu verschieden sind, bzw. nicht in einer beobachtbaren Beziehung zueinander stehen.

Diese Modelleinschränkung lässt sich reduzieren durch die Anwendung von korrelierten Frailty-Modellen, die nun nicht mehr von gemeinsamen Frailties für alle Mitglieder eines Clusters ausgehen.

Die Überprüfung auf unbeobachtete Heterogenität bei Shared Frailty erfolgt mit der Chi-Quadrat-Statistik als einseitiger Hypothesen-Test durch:

$$H_0 : \sigma^2 = 0 \text{ gegenüber } H_0 : \sigma^2 > 0$$

Sind die Werte für σ^2 sehr klein und nähern sie sich $H_0 : \sigma^2 = 0$, dann kann die Chi-Quadrat-Statistik in einigen Fällen das Testergebnis nicht klar voneinander abgrenzen.

Dann sollte bei einem positiven Nachweis von unbeobachteter Heterogenität dieses Ergebnis als eine reine Tendenz interpretiert werden (Wienke, 2006; Claeskens et al., 2008).

7.7 Alternative Frailty-Ansätze

7.7.1 Correlated-Frailty-Modelle

Bei den Correlated-Frailty-Modellen können die Frailties auch als relative Risiken definiert werden, diese müssen aber nicht zwingend identisch sein. Vielmehr ist das Ziel möglicher Korrelationen der Frailty-Eigenschaften für die einzelnen Individuen zu ermitteln. Dazu werden aber bivariate Verteilungsannahmen für die Frailties benötigt, welche z.B. durch drei unabhängige gammaverteilte Zufallsvariablen (Yashin et al. 1995a; Hens et al., 2009) quantifiziert werden.

Jede einzelne Zufallsvariable (vgl. Hens et al. 2009, S.2789) wird als additive Komponente $Y_i, i = 0, 1, 2$ beschrieben, welche jeweils für zwei Frailties

$Z_i = \sigma_i^2 (Y_0 + Y_i), i = 1, 2$ identisch ist. Die Korrelation zwischen den Frailties wird durch $\rho = k_0[(k_0 + k_1)(k_0 + k_2)]^{-1/2}$ definiert, wobei k_0, k_1, k_2 jeweils positive Werte annehmen, für $Y_i \sim \Gamma(k_i, 1) (i = 0, 1, 2)$

und $\sigma_i^2 = (k_0 + k_i)^{-1} (i = 1, 2)$ gilt.

Die Survivalfunktion kann dann für σ_1, σ_2 und ρ wie folgt formuliert werden:

$$S(t_1, t_2) = [S_1(t_1)]^{1-(\sigma_1/\sigma_2)\rho} [S_2(t_2)]^{1-(\sigma_2/\sigma_1)\rho} [S_1^{-\sigma_1^2}(t_1) + S_2^{-\sigma_2^2}(t_2) - 1]^{-\rho/\sigma_1\sigma_2} \quad (7.1)$$

Die Spezifikation der Zufallsvariablen muss nicht zwangsläufig mit einer Gamma-Verteilung erfolgen, sondern kann durch jede andere parametrische Form erfolgen.

7.7.2 Semiparametrische Frailty-Modelle mit EM-Algorithmus

Die Maximierung der Gamma-Frailty-Parameter kann auch semiparametrisch durch einen Expectation-Maximization (EM)-Algorithmus erfolgen. Die Iterationen des EM erfolgt zweistufig, mit jeweils einer Erwartungs- und Maximierungsprozedur (Duchateau und Jansen, 2008). In der ersten Stufe werden die Erwartungswerte für die unbeobachteten Frailties durch die vorhandenen direkt beobachtbaren Modellinformationen bestimmt.

Mit der zweiten Iterationsstufe werden die Erwartungswerte den wahren Werten für die Frailties angepasst, die Likelihood-Maximierung stützt sich wiederum auf die Erwartungswerte für die Frailties.

Diese methodische Vorgehensweise weist allerdings auch einige Schwächen auf. Die Modellanpassung durch einen EM-Algorithmus ist zeitlich aufwendiger und benötigt je nach Verfügbarkeit und Komplexität des Modells wesentlich mehr Ressourcen. Selbst bei sehr einfach strukturierten Modellen benötigt der EM-Algorithmus eine Vielzahl von Iterationen und kann daher in einigen Fällen die Modellanpassung nicht vollständig ausführen.

7.7.3 Die Penalized-Partial-Likelihood-Methode (PPL)

Alternativ zum EM-Algorithmus lässt sich durch eine Approximation mit der Penalized-Partial-Likelihood (PPL)-Methode (Therneau und Grambsch, 2000; Ripatti und Palmgren, 2000; Rodeau et al., 2003; Therneau et al., 2003) eine semiparametrische Schätzung durchführen. Charakteristisch für PPL ist die Einführung eines Penalty-Terms, der in ähnlicher Funktion eines Abzinsfaktors die Modellanpassung beschleunigt und mit Hilfe des Akaike-Information-Kriteriums (AIC) gültige Schätzungen für die Standardfehler liefert.

Die Schätzung erfolgt (vgl. Rondeau et al. 2003, S.142) semiparametrisch für die Parameter β, θ und einen geglätteten Baseline-Hazard $\lambda_0(t)$

$$pl(\lambda_0(\cdot), \beta, \theta) = l(\lambda_0(\cdot), \beta, \theta) - \sum_{h=1}^K \int_0^{\infty} (\lambda_{0h}'')^2(t) dt \quad (7.2)$$

mit $l(\lambda_0(\cdot), \beta, \theta)$ als Full Likelihood für die unbekannte Varianz θ und den Vektor für den Regressionskoeffizienten β .

Der zweite Teil der Gleichung bestimmt den Penalty-Term. Mit dem Glättungskoeffizienten K erfolgte die Gewährleistung eines Gleichgewichts zwischen der Modellanpassung und dem Glättungsgrad der geschätzten Hazardfunktion (Commenges et al., 1998). Die Schätzung des Penalized Likelihood wird mit einer nichtparametrische Regressionsfunktion, beispielsweise mit einer kubischen Spline-Funktion, realisiert.

7.7.4 Das Frailty-Konzept bei seltenen Ereignissen

Die Anwendung der Compound-Possion-Verteilung für die Frailty-Eigenschaften führt gerade bei relativ seltenen Ereignissen zu besonders reliablen Schätzergebnissen. Diese besondere Modellvoraussetzung erlaubt eine Betrachtung von Populationen, bei denen nur ein bestimmter Anteil von Individuen eine erhöhte Anfälligkeit gegenüber einer bestimmten Krankheit (Aalen und Tretli, 1999; Wienke et al., 2010) aufweist.

Beispielhaft kann hierbei die Inzidenz von Nasopharyngeal Carzinom (NPC) (Chan et al., 2005; Chang et al., 2006) angeführt werden, welches in Südostasien und Nordafrika häufiger auftritt als in den westlichen Industrieländer und sich spezifisch durch bimodale altersspezifische Inzidenzraten kennzeichnet.

Charakteristisch für diese Krebsform sind dabei die zwei beobachtbaren Peaks, einer im jungen Erwachsenenalter (15–24 Jahre) und der andere im höheren Erwachsenenalter (65–74 Jahre).

Haugen et al. (2009) haben daraufhin ein multiplikatives Frailty-Modell mit jeweils zwei Frailty-Termen für jeweils einen Peak und zwei Baseline-Raten entwickelt, um die Existenz eines zweimaligen Ansteigens der altersspezifischen Inzidenzen beweisen zu können.

$$h(t|Z_1, Z_2) = Z_1\lambda_1(t) + Z_2\lambda_2(t) \quad (7.3)$$

Die Besonderheit des Modells (vgl. Haugen et al. 2009, S.503) liegt in der Annahme einer Unabhängigkeit der beiden Frailty-Terme Z_1 und Z_2 , welche compound-poisson-verteilt sind. Dadurch ist es möglich, die Hazardraten (Weibull-verteilt) isoliert für jeden einzelnen Peak zu schätzen.

Ob diese Unabhängigkeits-Hypothese zwischen den beiden Frailty-Termen erfüllt ist, kann zum heutigen Kenntnisstand nicht belegt werden.

7.8 Mögliche Erklärungsfaktoren für unbeobachtete Heterogenität

Der Einfluss von unbeobachteter Heterogenität konnte zwar nicht anhand der verwendeten Teststatistik nachgewiesen werden, dennoch scheint es nicht ausgeschlossen zu sein, dass mögliche nicht beobachtbare Einflussgrößen die Modellschätzung beeinflussen könnten.

Unbeobachtete Heterogenität tritt immer dann auf, wenn bestimmte Faktoren und Variablen bei der Analyse unzureichend berücksichtigt wurden bzw. vollkommen unbekannt sind.

Insbesondere verhaltensspezifische Aspekte oder genetische Unterschiede können oftmals als unbeobachtete Komponente in vielen epidemiologischen Studien identifiziert werden, da sie oftmals aufgrund der vorliegenden Datengrundlage untererfasst sind.

7.8.1 Der Zusammenhang zwischen Rauchverhalten und Unfruchtbarkeit bei Männern

Schätzungsweise ein Drittel der gesamten Bevölkerung über dem 15. Lebensjahr sind aktive Raucher (El-Tonsy et al., 2009) und besitzen somit ein höheres Risiko für kardiovaskuläre Erkrankungen sowie die Bildung von malignen Neoplasien. Neben den bekannten negativen Einflüssen des Rauchens auf den Gesundheitszustand werden auch Störungen auf die reproduktiven Funktionen vermutet.

Ähnlich wie bei der Diskussion eines potentiellen Zusammenhangs von Fertilität und der postreproduktiven Sterblichkeit, wie in Kapitel 1.2 beschrieben, zeigen die einzelnen Studien sehr unterschiedliche Ergebnisse.

So konnte ein Großteil der Studien (Colagar, 2007; Soares, 2009; Zitzmann et al., 2004; Zavos et al., 1999) einen negativen Einfluss des Rauchens auf die Spermienkonzentration, Motilität und die Morphologie nachweisen.

In anderen wissenschaftlichen Arbeiten (Hoidas et al., 1985; Ozgur et al., 2005; Takase et al., 2004) konnte keine Assoziation zwischen dem Rauchverhalten und der Spermienqualität ermittelt werden.

Verantwortlich für die Folgeerscheinungen wie erektile Dysfunktionen, Herabsetzung der Spermiedichte und die Gesamtzahl sowie die Zerstörung der Spermien-DNA sollen die toxischen Nebenprodukte des Zigarettenrauchs sein (El-Tonsy et al., 2009). Bei Langzeitrauchern konnte somit auch eine deutlich höhere Konzentration von Cadmium und Oxidantien im Blut nachgewiesen werden (Al-Bader et al., 1999).

Scheinbar hat die Exposition gegenüber dem Nikotin und dessen Nebenprodukten einen schädigenden Einfluss auf die endothelvermittelte Vasomotorik (Zitzmann et al., 2004), welche für die Erektion des männlichen Schwellkörpers verantwortlich ist. Erektile Dysfunktionen können auch als Warnzeichen für eine Atherosklerose verstanden werden, weil diese oftmals komorbide mit anderen Folgeerkrankungen wie Diabetes auftreten.

In einer Querschnittstudie von 2010 Männern (Mirone et al., 2002) konnte für männliche Raucher und Ex-Raucher ein doppeltes Risiko von Anfälligkeiten gegenüber erektilen Störungen nachgewiesen werden.

Beeinträchtigend auf die Spermienkonzentration bzw. die Beweglichkeit der Spermien soll der Inhaltsstoff Cotinin (Abbauprodukt des Nikotins) wirken.

Die tägliche Anzahl der konsumierten Zigaretten und der damit einhergehenden, über die Jahre erhöhten Cotinin-Konzentration im Urin von Rauchern wirke sich negativ auf die Spermiedichte und die Gesamtzahl im Ejakulat auf (Zitzmann et al., 2004).

Neben dem Cotinin können relativ hohe Konzentrationen von Zink und Kupfer im Spermienplasma eine Verminderung der Spermienkonzentration und der Beweglichkeit der Spermien hervorrufen. Bei experimentellen Verpflanzungen von gesunden Spermien (Nichtrauchern) in das Seminalplasma von Rauchern zeigte sich, dass die Spermienbeweglichkeit deutlich abgenommen hat (Zavos et. al, 1998).

Andererseits konnte eine Verbesserung der Spermienparameter festgestellt werden, sobald die Spermien von Rauchern in das gesunde Seminalplasma umgesetzt wurden. Die Spermien-DNA kann auch durch den oxidativen Stress gestört werden. Bei einer übermäßigen Exposition von reaktiven oxygenen Spezies (ROS)⁷ im Seminalplasma wird die DNA angegriffen.

Die Anordnung der Basenpaare verändert sich, es kommt zu mehrfachen DNA-Crosslinks oder zur Zerstörung einzelner oder beider DNA-Stränge.

In dem Erbgut von Rauchern ist ein deutlich höherer Anteil von BPDE-I-dG-DNA nachweisbar, einer Verbindung aus einem diol-Epoxid, welches bei der Oxidation des

Benzopyren⁸ entsteht und sich mit dem Guanosin der DNA verbindet (Soares, 2009; Zenzen et al., 1999a).

⁷ Reactive oxygen species (ROS) sind Ionen bzw. sehr kleine Moleküle, die sowohl organische als auch anorganische Bestandteile enthalten: Sauerstoff-Ionen, freie Radikale und Peroxide. ROS treten bei allen Stoffwechselfunktionen in Verbindung mit Sauerstoff auf.

In Stresssituationen (zum Beispiel bei UV- oder Wärme-Exposition) steigt der ROS-Anteil an, sodass diese eventuelle Zellschädigung zur Folge hat.

⁸ Benzopyren ein aromatischer Kohlenwasserstoff, welcher hochmutagen ist.

Derartige Schädigungen der DNA sind schon in sehr frühen Embryonal-Stadien nachweisbar, wobei das Risiko einer Weitergabe eines DNA-Defektes durch die Eizellen im Vergleich mit der DNA-Läsion von rauchenden Vätern deutlich niedriger ist (Zenzes et al., 1999b).

7.8.2 Äußere Umweltfaktoren und männliche Unfruchtbarkeit

Bestimmte Umweltfaktoren und verhaltensspezifische Komponenten können Erkrankungen und Fehlbildungen des männlichen Genitaltraktes hervorrufen. Zu den verhaltensbedingten Faktoren zählen Alkohol- und Tabakkonsum, die Ernährung, Bekleidungsaspekte (wie z.B. das Tragen von zu engen Hosen) sowie eine übermäßige Exposition an Wärme und UV-Licht bei intensiven Solariums- und Saunabesuchen (Jakobi, 1999; Turek, 2008).

Chemische Substanzen und physikalische Einflüsse können die Hoden oder andere Organe schädigen und somit sämtliche testikuläre Funktionen und die Regulation der Sexualhormone betreffen. Solche Chemikalien können sogar die Wirkung körpereigener Hormone imitieren (Agonisten), in die Hormonsynthese direkt eingreifen oder ihnen sogar entgegenwirken (Antagonisten) (Sonnenschein und Soto, 1998). Diskutiert wird dabei auch der Zusammenhang zwischen einer Exposition gegenüber östrogenähnlichen Substanzen in der frühkindlichen Phase und einer Abnahme der Spermienqualität im Erwachsenenalter (Sharpe und Shakkebaek, 1993; Sonnenschein und Soto, 1998).

Dabei handelt es sich entweder um natürliche Inhaltstoffe (Phytoöstrogene⁹), die über die Nahrung aufgenommen werden, oder um synthetische Stoffe wie beispielsweise Pestizide und poly-chlorierte Biphenyle.

Andere Substanzen wie Fungizide, Vinclozolin und Procymidon kennzeichnen sich durch ihre antiandrogene Wirkung und können, in ähnlicher Weise, wie die östrogenähnlichen Chemikalien gewisse Störungen bei der sexuellen Differenzierung hervorrufen können, im Rahmen der Hodenabsenkung bleibende Schädigungen der Hoden und Nebenhoden hervorrufen (Gray, 1998).

Seit nunmehr zehn Jahren wird auch über ein Verbot der so genannten Weichmacher¹⁰ bei der PVC-Herstellung diskutiert. Gerade PVC-haltiges Kinderspielzeug gilt immer noch als sehr belastet und enthält oftmals zu hohe Konzentrationen dieser Substanzen, die nicht nur karzinogen sind, sondern sich auch negativ auf die Reproduktionsfähigkeit im späteren Erwachsenenalter auswirken können (Petersen, 2005).

Das DEHP (Di-(2-Ethylhexyl)-Phthalat) wurde von dem Centre for the Evaluation of Risk to Human Reproduction (CERHR) als „bedenklich für die menschliche Fortpflanzung“ eingestuft und betrifft insbesondere männliche Nachkommen (Eikmann und Herr, 2003; CERHR, 2006).

⁹ Phytoöstrogene ähneln strukturell sehr dem natürlichen Östrogen im menschlichen Körper.

¹⁰ Als Weichmacher können schwerflüchtige Ester, fette Öle oder Weichharze verwendet werden, welche Duroplasten bzw. Thermoplasten zugesetzt werden, um diese geschmeidiger und elastischer werden zu lassen.

7.8.3 Genetische Komponenten der männlichen Unfruchtbarkeit

Mutmaßlich scheinen in 30% aller Fälle von männlicher Unfruchtbarkeit genetische Störungen verantwortlich zu sein (Küpker et al., 2000). Die menschliche DNA besteht aus etwa 30.000 Genen (Ganten und Ruckpaul, 2008), die bestimmte lebensnotwendige Proteine verschlüsseln.

Tritt bereits an einer Stelle des DNA-Strangs eine fehlerhafte Basenpaarung auf, so kann sich diese nachhaltig auf die gesamte Proteinsynthese auswirken und bereitet somit schwere phänotypische Konsequenzen. Für die menschliche Fertilität sollen insgesamt 1.000 Gene verantwortlich sein. Nach gegenwärtigem Kenntnisstand wird eine „Störung der Spermiogenese ursächlich bei einer chromosalen Aneuploidie bei Mutationen in putativen Spermatogenesegenen“ vermutet (Küpker et al., 2000, S. 79).

Aus historischer Sicht konnte ein systematischer Zusammenhang zwischen den chromosomalen Anomalien und Infertilität bereits in den späten 50er Jahren des 20. Jahrhunderts nachgewiesen werden (Ferguson-Smith et al., 1957). Etwa zur gleichen Zeit wurde auch der Karyotyp 47, XXY, bestimmt, welcher verantwortlich für die Subfertilität bei Männern ist (Jacobs und Streng, 1959).

Die am häufigsten diagnostizierte chromosomale Anomalie ist das Klinefelter-Syndrom (ca. 5%) mit einer 44-fach höheren Prävalenz gegenüber der fertilen männlichen Vergleichsgruppe (Küpker et al., 2000).

KS folgt aus einer „non-disjunction“ des X-Chromosoms und wird in 93% aller Fälle durch den Karyotyp 47, XXY, charakterisiert (Baselmaier und Tariverdian, 2007; Bielanska et al., 2000; Hynh et al., 2002).

Das Krankheitsbild ist gekennzeichnet durch ein geringes Hodenvolumen und eine häufige Azoospermie. Der Testosteronspiegel ist sehr niedrig, dagegen ist die Konzentration von Östradiol im Blut erhöht.

Bei einigen KS-Patienten können wenige bewegliche Spermatozoen im Ejakulat aufgefunden werden, sodass sich oftmals nur eine testikuläre Spermatozoenextraktion (TESE) als therapeutische Option empfiehlt.

Das Kallmann-Syndrom ist die häufigste Form eines Gen-Defektes, die das X-Chromosom betrifft. Es handelt sich hierbei um eine rezessiv vererbare Mutation des KALIG-1-Gens auf Xp 22.3 (Franco et al., 1991). Bei Männern ist das Kallmann-Syndrom karyotypisch häufiger präsent als bei Frauen und gilt als eine Unterform des Hypogonadismus. Als Begleit leiden treten unter anderem auch Anosomie (Verlust des Geschmackssinns), Gehörlosigkeit, Lippen-Kiefer-Gaumensplate und Farbenblindheit auf (EAU, 2004).

Eine eher seltene Form einer Gen-Mutation des X-Chromosoms ist die Androgenresistenz, auch bekannt als Reifenstein-Syndrom, ein Defekt des Androgen-Rezeptor-Gens Xq11-12. Die Patienten weisen oftmals eine hohe Konzentration an Testosteron und niedrigere Spermienzahlen auf, genitale Anomalien sind dagegen selten.

Neben den genetischen Störungen, die das X-Chromosom betreffen, können auch Mikrodeletionen des Y-Chromosoms einen Einfluss auf die Spermiogenese haben.

Als eine Verbindung von Mikrodeletionen auf dem langen Arm (Yq₁₁) des Y-Chromosoms¹¹ und Yq11 befinden sich drei nicht überlappende Regionen AZF-a-b-c (Vogt et al. 1996; EAU, 2004), von denen AZF-c als jener Lokus identifiziert wurde, auf dem sich das DAZ- Gen (Deleted in Azoospermia) befinden soll.

Der Nachweis einer Deletion des DAZ-Gens konnte aber auch eine gewisse Heterogenität im Phänotyp der Patienten aufzeigen, welcher zwischen Azoospermie, oligozoospermen, in wenigen Fällen auch Fertilität variiert (Küpker, 2000).

Die Ergebnisse einer Mapping-Studie (Silber, 2003) folgen der Argumentation, dass eine Deletion auf dem Y-Chromosom nicht unbedingt die Fertilitätschancen von azoospermen und oligozoospermen Patienten minimiere, da diese chromosomale Anomalie nur in sehr wenigen Fällen (10% bei Azoospermie) nachgewiesen werden konnten. Bei nichtobstruktiver Azoospermie und Oligozoospermie zeigen sich markante irreversible Störungen der Spermproduktion in den Hoden können in 15 % mit dieser Form der Azoospermie auf chromosomale Anomalien zurückgeführt werden.

¹¹ Das Y-Chromosom lässt sich in drei Bestandteile aufgliedern: Euchromatin auf dem kurzen Arm (Yp₁₁), Euchromatin auf dem langen Arm (Yq₁₁) und Heterochromatin auf dem langen Arm (Yq₁₂)

In 13% der Fälle besteht eine gewisse Chance für Mikrodeletionen des Y-Chromosoms (Turek, 2008). Das Risiko für chromosomale Anomalien ist bei oligozoospermen Patienten mit nur 2% deutlich geringer, ebenso wie das Potential einer möglichen Beschädigung des Y-Chromosoms (Turek, 2008).

Die Mikrodeletionen des Y-Chromosoms können nur vom Vater ausgehend an die männlichen Nachkommen vererbt werden, potentielle Mutationen des X-Chromosoms werden dagegen nur an die Töchter weitergegeben.

Wie und in welchem Ausmaß sich diese genetischen Störungen auch auf die Fertilität auswirken, kann beim heutigen Kenntnisstand noch nicht eindeutig geklärt werden, da bisher nur sehr wenige Fälle existieren, bei denen eine eindeutige Vererbung bestimmt werden konnte.

Tabelle 28: Reference for Genetic Risks associated with Common Male Infertility (Turek, 2008)

Diagnosis	Chromosome Abnormality Risk	Y Chromosome Microdeletion Risk	CF Mutation/ST Allele Risk
Oligozoospermia (<5 million sperm/ml)	2-7%	6-8%	same as general population
Nonobstructive Azoospermia	15%	13%	same as general population
Congenital unilateral absence of vas deferens CUAVD 15%	not increased	not increased	up to 50%
Congenital bilateral absence of vas deferens CBAVD	not increased	not increased	up to 80%

Die congenitale Absenz des vas deferens umfasst alle physiologischen Störungen, die den Transport der Spermien über die Samenstränge betreffen, die mit einer angeborenen „natürlichen Vasektomie“ gleichzusetzen sind.

Die congenitale Absenz des vas deferens umfasst alle physiologischen Störungen, die den Transport der Spermien über die Samenstränge betreffen, die mit einer angeborenen „natürlichen Vasektomie“ gleichzusetzen sind. Diese Schädigung des vas deferens ist mutmaßlich auf eine Mutation der cystischen Fibrose-Gene (z.B. eine Deletion des Phenylalaninrestes an Position 508 des CFTR-Genes) zurückzuführen und tritt, wie der Tabelle 55 zu entnehmen ist, häufiger beidseitig auf (Cuppens and Cassiman, 2004; Guggino und Stanton, 2006).

Tabelle 29: Mit Infertilität verbundene genetische Störungen
(European Association of Andrology)

Genetischen Störung	Phänotyp	Art der Vererbung
Prader-Willi-Syndrom	Fettleibigkeit, psychische Störungen	Defekt auf 15q12 weitergegeben durch den Vater
Bardet-Biedle-Syndrom	Fettleibigkeit, psychische Störungen Retinitis pigmentosa, Polydactylie	autosomal rezessives Syndrom, q21
Cerebellar ataxia und hypogonadrophischer Hypogonadismus	Eunuchoidismus, Gang- und Sprach- störungen	autosomal rezessiv
Noonan's Syndrom	Wachstumsverzögerung und Minderwuchs, verbreitete Nackenhaut Beeinträchtigungen für Herz- und Lunge Kryptorchismus	autosomal dominant
Myotische Dystrophie	Muskel-Schwund, Sehstörungen (grauer Star) Atrophie	autosomal dominant, 19q13.3
Dominante polyzystische Nierenentzündung	Nieren-Zysten, Epididymal-Zyste	autosomal dominant, 19q13.3

Darüber hinaus erfolgte durch die European Association of Andrology (Dohle et al., 2004, 2007) eine Auflistung weiterer genetischer Störungen, die zum einen eher selten sind und zum anderen mit Ausnahme des hypogonadrophischen Hypogonadismus eher zeitrangig in Verbindung mit Fertilitätsstörungen gebracht werden.

8. Zusammenfassung

Die vorliegende wissenschaftliche Arbeit hat als Zielsetzung zu klären, ob die Variable Fertilitätsstatus bei Männern ein geeigneter Indikator für mögliche Sterblichkeitsunterschiede ist. Können entsprechende Disparitäten in der postreproduktiven Mortalität durch diese Einflussgröße direkt miteinander in Verbindung gebracht werden, oder sind andere Faktoren wie z.B. unbeobachtete Heterogenität für diese scheinbaren Unterschiede verantwortlich.

Der Einfluss von unbeobachteter Heterogenität lässt sich durch eine positive latente Zufallsvariable, die Frailty Z , beschreiben. In vielen wissenschaftlichen Studien zeichnen sich die Untersuchungspersonen durch eine Diversität bezüglich ihrer individuellen Merkmale aus. Personen können daher sehr unterschiedliche Frailty-Eigenschaften haben und sich daher in ihrer individuellen Anfälligkeit gegenüber bestimmten Ereignissen wie Krankheit oder Tod unterscheiden.

Kann eine entsprechende Varianz für die Frailties innerhalb der Untersuchungspopulation nachgewiesen werden, so könnten möglicherweise die Parameterschätzungen, insbesondere bei den Proportional-Hazards-Modellen, verzerrt sein bzw. die Interpretation dieser Ergebnisse unpräzise werden, sobald die Proportional-Hazards-Annahme nicht mehr aufrechterhalten werden kann.

Als Untersuchungsgrundlage für diese Studie werden Spermioqrammdaten von Patienten der Andrologischen Abteilung der Universitätsklinik Marburg verwendet, welche sich über den Zeitraum von 1949 bis 1998 überwiegend bei Nichterfüllung eines Kinderwunsches einer Untersuchung ihrer Spermienparameter unterzogen haben. Die Überprüfung des Vitalstatus erfolgte für alle Patienten, die vor dem 01.01.1941 geboren wurden, zum Stichtag 31.12.2006 über das Einwohnermeldeamt, das hessische Zentralarchiv einer gesetzlichen Krankenversicherung.

Insgesamt konnten 1408 ehemalige Patienten über mindestens eines dieser Suchsysteme identifiziert werden, die übrigen 888 Fälle sind von Lost-to-Follow-Up betroffen und können für die Analyse nicht berücksichtigt werden. Die systematische Abgrenzung der unterschiedlichen Fertilitätszustände fertil und subfertil orientiert sich an den WHO-Empfehlungen (2010).

Die Subfertilen könnten zusätzlich nach Oligozoo- und Azoospermen unterschieden werden, wobei Letztere keine Spermatozoen im Ejakulat aufweisen. Jene Fälle mit eher unauffälligen Spermienparameter werden als normozoosperm klassifiziert.

Die Punktschätzungen für die mittlere Lebensdauer auf Basis der Greenwood-Formel ergaben für die subfertilen Fälle eine kürzere Lebensdauer von 2,4 Jahren gegenüber der fertilen Vergleichsgruppe.

Die durchschnittliche Lebensdauer der Normozoospermen gegenüber den oligozoospermen Patienten hat sich ebenfalls um 2,4 Jahre verlängert, der Unterschied der mittleren Lebenserwartung im Vergleich von normozoospermen und azoospermen Patienten beträgt 2,3 Jahre.

Die nichtparametrischen Schätzverfahren zeigten sowohl rechnerisch – die Überprüfung erfolgte mit Log-rank, Wilcoxon, Tarone-Ware und Peto-Peto-Prentice – als auch graphisch keine signifikanten Sterblichkeitsunterschiede für subfertile und fertile Fälle auf. Auch bei der Betrachtung der drei Fertilitätsklassifikationen azoosperm, oligozoosperm und normozoosperm konnten ebenfalls keine signifikanten Unterschiede festgestellt werden.

Da die nichtparametrischen Verfahren zur Kontrolle der Confounder-Effekte keine weiteren Kovariablen zulassen, erfolgten zunächst Schätzungen mit den Proportional-Hazards-Modellen nach Cox. Es konnten nur schwache signifikante Sterblichkeitsunterschiede von fertilen und subfertilen Männern ermittelt werden, dagegen unterscheiden sich normozoosperme, oligozoosperme und azoosperme Fälle nicht signifikant in ihrer Hazardneigung.

Die numerische Teststatistik für die Proportional-Hazards-Annahme mit den Schoenfeld-Residuen ergab keine Verletzung der Proportional-Hazards-Bedingung, folglich müssen die Parameterschätzungen nach Cox nicht verworfen werden.

Die Interpretation der graphischen Überprüfung der PH-Annahme mit den Log(log)-Hazards-Verläufen über die Beobachtungszeit war dagegen etwas schwieriger, da gerade bei zeitlich-kontinuierlichen Variablen und bei kategorialen Merkmalen nicht ausgeschlossen werden kann, dass sich die Hazard- und Survivalkurven überlagern und dann nicht mehr voneinander unterschieden werden können.

Die parametrischen Regressionsmodelle mit Unshared Frailty konnten nur bei der Gompertz(Gamma)-Spezifikation fertilitätsabhängige Unterschiede in der Hazardneigung bestimmen, diese gelten aber nur bei der Betrachtung des dichotomisierten Fertilitätsstatus (subfertil und fertil). Es konnten keine statistischen Einflüsse durch die anderen Kovariablen ermittelt werden. Der Nachweis der Frailty-Varianz, ergab, dass die Nullhypothese von $H_0 : \theta = 0$ nicht abgelehnt werden konnte. Die beste Modellanpassung liefert nach AIC das Gompertz(Gamma)-Modell.

Aufgrund der dargestellten Ergebnisse ist allgemein davon auszugehen, dass die Spermienkonzentration kein besonders guter Indikator ist um Unterschiede in der Sterblichkeit zu ermitteln. Ebenso konnte für keine der verwendeten Kovariablen Geburtsjahr und Wohnort ein Effekt festgestellt werden.

9. Fazit und Ausblick

Abschließend erfolgen ein Resümee und eine kontextuelle Einordnung der Studienergebnisse für die Relevanz in der Mortalitäts- und Fertilitätsforschung.

Rückblickend auf die Ausgangsfrage eines postulierten möglichen Zusammenhangs von Fertilität und postreproduktiver Sterblichkeit lässt sich schlussfolgern, dass nur sehr wenige Anhaltspunkte für eine statistische Assoziation dieser beider Effektgrößen bestehen. Es existieren bereits einige empirische Untersuchungen, die eine zeit- und altersabhängige Veränderung der Spermienqualität beweisen konnten (Eskenazi et al., 2001; 2003)

So ist mit einer Reduktion des Spermienvolumens von 0,03ml pro Altersjahr zu rechnen. Auch die Motilität sinkt um 0,7 % pro Jahr, wobei die jährliche Abnahme der progressiven beweglichen Spermien mit 4,7 % doch sehr hoch zu sein scheint. Ebenso sollen sich die Anteile an beweglichen Spermien bei einem 50-jährigen Mann um 28% gegenüber dem Niveauwert eines 30-jährigen Mannes reduzieren. Aus der altersspezifischen Abnahme der Spermienqualität könnte ein altersbedingtes Ansteigen der Mortalität folgen, deshalb scheint die Frage eines statistischen Zusammenhangs weiterhin ungeklärt zu bleiben.

In den zitierten Arbeiten zur Lebenszeitmortalität von Männern mit unterschiedlichen Fertilitätswuständen konnte für subfertile Männer ein 1,6-fach höheres Sterberisiko nachgewiesen werden (Groos, 2005; Groos et al., 2006). Dieser Effekt erwies sich bei Erweiterung des Datensatzes unter der Berücksichtigung verlängerter Beobachtungszeiten als instabil.

Andererseits konnte aber in einem Papier von Jensen et al. (2009), erschienen im *American Journal of Epidemiology*, mit einer weitaus größeren Datenbasis ($n = 43.277$) eine geringere Sterblichkeit dänischer Männer bei positiven Spermienparametern nachgewiesen werden. Allerdings sind die methodische Vorgehensweise und die verwendeten Auswertungsverfahren in diesem Papier nur unzureichend dokumentiert, sodass einzelne Ergebnisse dieser Studie kritisch beurteilt werden sollten.

Darüber hinaus heißt es in Jensen et al. (2009, S. 561): „... to our knowledge, only one study has investigated the effect of poor semen quality on subsequent mortality, and this study suffered from serious methodological problems and was not age adjusted“, womit die hier in Marburg entstandene Studie von Sabine Groos (2005) gemeint ist.

Diese scheinbar gerechtfertigte Kritik lässt sich teilweise zurückweisen, da sich doch ein Großteil der Ergebnisse auch bei Anwendung komplexerer Auswertungsverfahren replizieren lässt. Dennoch bleibt die Frage offen, inwiefern die Fertilität eines Mannes eine Bedeutung für dessen postreproduktive Sterblichkeit hat.

Der vermeintlich geringere Einfluss einer Vaterschaft auf die Sterblichkeit lässt sich erneut mit evolutionsbiologischen Argumenten untermauern. Das klassische Frauenbild war geprägt durch die Erfüllung der weiblichen Lebensaufgabe mit einer Maximierung der Nachkommenschaft, die zum einen für die Haushaltsproduktion, aber auch für die Altersversorgung wichtig war. Die Reproduktionskosten von Frauen und Männern sind insbesondere in schwierigen Lebenssituationen sehr unterschiedlich (Gagnon, 2005).

So mussten Frauen im Krankheitsfall oder beim Tod des Ehepartners häufig die schweren körperlichen Arbeiten übernehmen, die das eigene Überleben zunächst sicherten.

Gleichzeitig mussten sie sich dann auch um die Versorgung und Erziehung des Nachwuchses kümmern.

Dieses lässt auch auf eine stärkere emotionale Bindung von Frauen an die Kinder schließen.

Ebenso könnte dann auch die psychische Belastung bei Nicht-erfüllung des Kinderwunsches bei Frauen wesentlich häufiger vertreten sein.

In einer empirischen Studie zur geschlechtsspezifischen Bewertung individueller psychischer Belastungen bei Infertilitätsdiagnose konnte gezeigt werden, dass bei einer ursächlichen Kinderlosigkeit, die beide Partner betrifft, Frauen über eine höhere Unzufriedenheit in der Ehe und Sexualität berichteten als ihre Partner (Lee et al., 2001).

Dabei sind die stärkeren Belastungsempfindungen von Frauen zum einem auf die oftmals sehr einseitige Intensivierung der Diagnose- und Behandlungsstrategien auf der weiblichen Seite liegende Ursache von Infertilität zu finden, zum anderen sind Frauen emotionaler und können körperliche und psychische Belastungen eher wahrnehmen und darüber berichten.

Insbesondere bei Männern ist die Bewältigung der individuellen Belastungen bei unerfülltem Kinderwunsch davon, abhängig inwiefern dieser väterliche Identitätsverlust durch ein stärkeres Engagement in anderen Lebensbereichen, z.B. im Beruf, „die Kränkung durch die Kinderlosigkeit“ kompensieren kann (Könnecke und Wilke, 2002).

Das immer noch sehr verbreitete traditionelle Rollenverständnis des „male bread-winner“ verdeutlicht, dass sich Männer nach wie vor eher über Beruf und Einkommen definieren und dadurch weniger Zeit für Betreuungs- und Erziehungsaufgaben investieren können (Schmitt, 2004; Schmidt und Winkelmann, 2005).

Bei einer oftmals unterstellten fehlenden emotionalen Bindung zum eigenen Nachwuchs sind einige Männer wahrscheinlich eher bereit vollständig auf die Vaterrolle zu verzichten. Dann könnte aber die Bedeutung von Partnerschaft und Freunden, die dann als Kompensationsgut für ungewollte Kinderlosigkeit genutzt werden, wichtiger werden.

In ihrer abschließenden Bemerkung zur individuellen Bedeutung von ungewollter Kinderlosigkeit äußern sich Könecke und Wilke wie folgt: „Ob jemand Kinder hat oder nicht, ist Bestandteil der eigenen Lebensgeschichte und damit, nach erfolgreicher Integration, Teil der eigenen Identität“ (Könecke und Wilke, 2002, S. 73).

Insgesamt lässt sich festhalten, dass die thematische Auseinandersetzung mit Männergesundheit und einer spezifischen Sterblichkeit in vielen Forschungsgebieten noch unterrepräsentiert ist.

Die Fertilitätsforschung hatte sich ebenfalls über viele Jahre nur auf die weibliche Perspektive konzentriert, was nicht immer nur an fehlendem Interesse für männerspezifische Themenfelder lag, sondern einfach auf Datenerhebungsprobleme zurückzuführen ist (Schmitt, 2005).

Im Mikrozensus beispielsweise wird zum Zeitpunkt der Erhebung nur die aktuelle Haushaltszusammensetzung erfasst. Eine Unterscheidung nach sozialer bzw. leiblicher Elternschaft erfolgt nicht. Bei getrennt lebenden Eltern teilen mit alleinigem Sorgerechtsanspruch der Ehefrau werden die Kinder dem Haushalt der Mutter zugeordnet. Die Zuordnung der biologischen Vaterschaft ist dann im Trennungsfall nicht mehr möglich.

In diesem Zusammenhang könnte eine zweigeschlechtliche Untersuchung dieses Problem minimieren. An der menschlichen Reproduktion sind beide Geschlechtspartner in gleicher Weise beteiligt.

Bei einer einseitigen Diagnose von Sterilität bei nur einem Partner kann nicht ausgeschlossen werden, dass ein Paar trotzdem eine Schwangerschaft erzielt. Ebenso kann ein Paar kinderlos bleiben, obwohl bei beiden Partnern keine physischen Anzeichen von Infertilität bestehen.

Daher sollte der Begriff Fertilität in diesem Kontext eher als Potential verstanden werden, zumal die Aussagekräftigkeit eines Spermioграмms für die Fertilität eines Paares nur bei ca. 28 Prozent liegt (Krause, 1993).

Immer noch ist die Datenlage in Bezug auf Längsschnittuntersuchungen individueller Gesundheitsdaten doch sehr überschaubar. Nicht nur langfristige Zusammenhänge zwischen den Einflussgrößen, sondern auch zeitliche Variabilität lassen sich mit longitudinalen Studien überprüfen.

Die vorliegende Studie leistet trotz eines feststellbaren nichtsignifikanten Zusammenhangs von Fertilität und postreproduktiver Sterblichkeit einen wichtigen Beitrag zu diesem wissenschaftlichen Feld.

Aalen, O.O. (1988): Heterogeneity in survival analysis. *Statistics in Medicine* 7, pp. 1121-1137.

Aalen, O.O. (1992): Modelling heterogeneity in survival analysis by the compound Poisson distribution. *Annals of Applied Probability* 4, pp. 951-972.

Aalen, O.O., Tretli, S. (1999): Analyzing incidence of testis cancer by means of a frailty model. *Cancer Cause and Control*, 10(4), pp.1569-88.

Abbring, J.H., Van den Berg, G.J. (2006): The Unobserved Heterogeneity Distribution in Duration Analysis. *Biometrika*, 94 (1), pp.87-99.

Abramowitz, M., Stegun, I.A. (1972): *Handbook of Mathematical Functions*. Dover, New York.

Al-Bader, A., Omu, A., Dashti, H. (1999): Chronic cadmium toxicity to sperm of heavy cigarette smokers: immunomodulation by zinc. *Archives of andrology*, 43 pp. 135-40.

Andersen, P.K. (1982): Testing Goodness of Fit of Cox-Regression and Life Model. *Biometrics*, 38, 67-77. Correction (1984): 40, p. 1217.

Andersen, P.K., Borgan, O., Gill, R.D., Keiding, N. (1996): Statistical Models Based on Counting Processes. Springer-Verlag, New York.

Anzures-Carera, J., Hutton, J.L. (2010): Competing risks, left truncation and late entry effect in A-Bomb survivors cohort. *Journal of Applied Statistics*, Vol 37, No.5, pp.821-831.

Aylor, D.E. (1990): The role of intermittent wind in the dispersal of fungal pathogens. *Annual Review of Phytopathology* 28, pp.73-92.

Baselmaier, W., Tariverdian, G. (2007): Humangenetik. 4. neue bearbeitete Auflage, Springer, Heidelberg, Berlin.

Baskin, L.S. (2000): Hypospadie and urethral development. *Journal for Urology* 163, No.3, pp. 951-6.

Berkman, L.F., Syme, S.L. (1979): Social networks, host resistance, and mortality: A nine-year follow-up study of alameda county residents. *American Journal of Epidemiology*, 109, pp. 186-204.

Bielanska, M., Tan, S.L., Ao, A. (2000): Fluorescence in-situ hybridization of sex chromosomes in spermatozoa and spare preimplantation embryos of a Klinefelter 46, XY/47,XXY male. *Human Reproduction*, 15, pp.440-444.

Butt, Z., Haberman, S. (2004): Application of Frailty-Based Mortality Models Using Generalized Linear Models. *Astin Bulletin*, Vol.34, No.1, pp.175-197.

Brinkworth, M.H., Handelsman, D.J. (2009): Umwelt- und arbeitsplatzbedingte Einflüsse auf die männliche Fertilität In: Nieschlag, E., Behre, M., Nieschlag, S. *Andrologie*, 3. Auflage, Grundlagen und Klinik der reproduktiven Gesundheit des Mannes. Springer Berlin/Heidelberg.

Bruckert, E. (1991): Wie häufig ist ungewollte Kinderlosigkeit? *Andrologia* 23, pp.245-250.

Carroll, K.J. (2003): On the use and utility of the Weibull modell in the analysis of survival data. *Controlled Clinical Trials*, Vol.24, No.6, pp.682-701.

Canales, B.K., Zapzalka, D.M., Cesar, J.R., Carey, P. und andere (2005): Prevalence and effect of varicuceles in an elderly population. *Urology*, 66, pp.627-631.

Chan, J.K., Bray, F., McCarron P., Foo, W., Lee, A.W.M. und andere (2005): Nasopharyngeal carcinoma. In: Barnes, E.L., Eveson, J.W., Reichart, P., Sidransky, D.(editors). *Pathology and Genetics of Head and Neck Tumors*.

Kleihues, P., Sohin, L.H.(series editor). World Health Organisation Classification of Tumors, Lyon, France: IARC Press, pp.85-97.

Chang, E.T., Adami, H.-O. (2006): The enigmatic epidemiology of nasopharyngeal carcinoma. *Cancer Epidemiology, Biomarkers and Prevention*, 15, pp. 1765-1777.

Chuang, S.K., Cai, T., Douglass, C.W., Wei, L.J. und Dodson, T.B. (2005): Frailty Approach for the Analysis of Clusterd Failure Time Observations in Dental Research. *Journal of Dental Research*, 84, p. 54.

Choi, J.H., Park, H.I. (2007): An nonparametric test for diagnosis of the proportional hazard assumption. *Statistical Papers*, 48, pp.467-477.

Claeskens, G., Nguti, R., Janssen, P. (2008): One-side test in shared frailty models. *Test*, 17, pp.69-82.

Clayton, D. (1978): A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, 65 (1), p.141.

Clebens, M.A., Gould, W.M., Gutierrez R.G., Marchenko, Y.U. An Introduction to Survival Analysis Using Stata. 2nd Edition, Stata Press Publication StataCorp LP , College Station Texas, 2008.

Coale, A.J., Kisker, E.E. (1986): Mortality Crossover: Reality or Bad Data? Population Studies, 40, pp.389-401.

Cohen, S. (2004): Social Relationship and Health. American Psychologist, Vol. 59(8) pp. 676-684.

Commenges, D., Letenneur, L., Joly, P., Aliouh, A., Dartigues, J.F.(1998): Modelling age-specific risk: Application to dementia. Statistics in Medicine, 17, pp.1973-1988.

Colagar, A.H., Jorsaraee, G.A., Marzony, E.T. (2007): Cigarette Smoking and the Risk of Male Infertility. Pakistan Journal of Biological Sciences, 10 (21), pp. 3870-3874.

Cox, D.R.(1972): "Regression models and life tables with discussion". Journal of the Royal Statistical Society, Series B, 34, pp.187-220.

Cuppens, H., Cassiman, J.J. (2004): "CFTR mutations and polymorphism in male infertility". International Journal of Andrology, 27, pp. 251-256.

Doblhammer, G., Oeppen, J. (2003): Reproduction and Longevity among the British peerage: The effect of frailty and health selection. Biological Sciences 270, pp. 1541-1547.

dos Santos, D.M., Davies, R.B., Francis, B. (1995): Nonparametric hazard versus nonparametric frailty distribution in modelling recurrence of breast cancer. *Journal of Statistical Planning and Inference*, 47, pp.111-127.

Dupre M.E., Franzese, A.T., Parrado, E.A. (2006): Religios Attendance and Mortality: Implications for the Black-White Mortality Crossover. *Demography*, 43 (1), pp.141-164.

Duchateau L., Janssen P. (2008): *The Frailty Model*. Statistics for Biology and Health. Springer Science + Business Media, Inc. New York.

Drefal, S. (2010): How does the age gap between partners affect their survival? *Demography*, 47 (2), pp.313-326.

Dribe M. (2004): Long-term effects of childbearing on mortality: evidence from pre-industrial Sweden. *Popul stud* 58, pp. 297-310.

Eikmann, T., Heer, C. (2003): Phthalate- eine bisher unterschätzte Belastung für die Bevölkerung. *Umweltmedizinische Forschung in der Praxis* 8 (1), pp.3-5.

Elo, I.T., Preston S.H. (1997): Racial and Ethnic Differences in Mortality at Older Ages. Chapter 2, pp. 10-42 In Martin & Soldo (Eds) *Racial and Ethnic*

Differences in the Health of Older Americans. Washington D.C.: National Academy Press.

El-Tonsy, H., Abdel-Raof, H., Hossam El-Din W., El-Fakahany, H., Gabr M., Othman A, Medhat, W. (2009): Smoking Arguments the Detrimental Effect of Varicocele in Infertile patients through Increased Level of Nitric Oxide in Spermatic Vein Blood. PALD Vol.20, No.1, pp.89-96.

Eskenazi, B., Wyrobek, A.J., Slotter, E., Kidd, S.A., Moore, L., Young, S. and Moore, D. (2003): The association of age and semen quality in healthy men. Human Reproduction Vol.18, No.2, pp.447-454.

Fathalla, M.F. (1992): Reproductive health: a global overview. Early Human Development 29, 1-3, pp.35-42.

Ferguson-Smith, M.A., Lennox, B., Mack W.S. et al. (1957): Klinefelter's syndrome: frequency and testicular morphology in relation to nuclear sex. Lancet II: 167.

Fleming, T.R. and Harrington, D.P. (1991) Counting Process Analysis. Wiley, New York.

Ficarra, V. et al (2006): Treatment of Varicocele in subfertile men: The Cochrane review A contrary opinion. European Urology, 49, pp. 258-263.

Finkelstein, M. (2007): Shocks in homogenous and heterogeneous populations. *Reliability Engineering and System Safety*, 92, pp. 569-574.

Finkelstein, M., Esaulova, V. (2006): Bounds for the failure rate in heterogeneous populations. *South African Statistical Journal*, 41, pp. 21-36.

Franco, B., Guioli, S., Pragliola, A. et al. (1991): A gene deleted in Kallmann's syndrome shares homology with neural cell adhesion and axonal pathfinding molecules. *Nature* 353, pp.529-536.

Gagnon, A, Mazan, R, Desjardins B, Smith K.R. (2008): Post-reproductive longevity in a natural fertility population. In: Bengtsson T, Mineau G, editors. *Kinship and Demographic Behavior in the Past*: Springer/ IUSSP. pp 225-241.

Ganten, D., Ruckpaul, K. (2008): *Grundlagen der molekularen Medizin*. 3. überarbeitete und erweiterte Auflage, Springer Medizin Verlag Heidelberg.

Giles, L.C., Glonek, G.F.V., Luszcz, M.A., Andrews G.R. (2005): Effect of social networks on 10 year survival in very old Australians: The Australian longitudinal study of aging. *Journal of Epidemiology and Community Health*, 59, pp. 574-579.

Guggino, W.B., Stanton, B.A. (2006): "New insights into cystic fibrosis: molecular switches that regulate CFTR", *Nature Reviews Molecular Cell Biology*, 7, pp. 426-436.

Grambsch, P.M., Therneau, T.M. (1994): Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, 81, 3, pp.515-526.

Gray, L.E. Jr. (1998): Xenoendocrine disruptors: laboratory studie on male reproductive effects. *Toxicology Letters*, 102-103, pp. 331-335

Groos, S., Krause, W., Mueller, U. (2006): Men with Subnormal Sperm Counts Live Shorter Lives. *Social Biology, Biannual Journal of the study of Social Biology*, Vol. 53, No. 1-2., pp.46-60.

Grundy, E. and Kravdal, Ø. (2008): Reproductive history and mortality in late middle age among Norwegian men and women. *American Journal of Epidemiology*, 167, pp. 271-279.

Guzick et al. (2001): Sperm morphology, motility and concentration in fertile and infertile men. *The New England Journal of Medicine*, 345, pp.1388-93.

Hall, W.J., Wellner, J.A. (1980): Confidence intervals and confidences bands for the cumulative hazard rate function and their small sample properities. *Biometrika*, 67, pp.133-143.

Haugen, M., Bray, F., Grotmol, T., Tretli, S., Aalen, O.O., Moger, T. (2009): Frailty modelling of bimodal age-incidence curves of nasopharyngeal carcinoma in low-risk populations. *Biostatistics*, 10, 3, pp.501-514.

Hautmann, R., Huland, H. (2006): *Urologie*. 3. überarbeitete Auflage, Springer Medizin Verlag Heidelberg.

Hause, R., Sokol, R. (2008): Science linking environmental contaminant exposures with fertility and reproductive health impacts in the adult male. *Fertility and Sterility*, 89, pp.59-65.

Henderson, R., Oman, P. (1999): Effect of Frailty on marginal regression estimates in survival analysis. *Journal of the Royal Statistical Society Series B (Statistical Methodology)*, 61 (2), pp. 367-379.

Hens, N., Wienke, A., Aerts, M., Molenberghs, G. (2009): The correlated and shared frailty model for bivariate current status data: An illustration for cross-sectional serological data. *Statistics in Medicine* 28, pp.2785-2800.

Hess, K.R. (1995): Graphical methods for assessing violations of the proportional hazards assumption in Cox Regression. *Statistics in Medicine*, 14, pp.1707-1723.

Hill, M.E., Preston S.H., Rosenwaike, I. (2000): Age Reporting among White Americans Ages 85+: Results of a Record Linkage Study. *Demography* 37, pp.175-186

Hoidas, S., Williams, A., Tocher, J., Hargreave, T. (1985): Scoring sperm morphology from fertile and infertile cigarette smokers using the scanning electron-microscope and image analysis. *Fertility and Sterility*, 4, 3, pp.596-598.

House, J.S., Kahn, R.L. (1985): Measures and concepts of social support. In: S. Cohen & S.L. Syme (Eds.), *Social support and health* pp.83-108. New York: Academic Press.

Hougaard, P. (1984): Life table methods for heterogenous populations. *Biometrika*, 71, pp. 75-83

Hougaard, P. (1986a): Survival models for heterogeneous populations derived from stable distributions. *Biometrika* 73, pp. 387-396.

Hougaard, P. (1986b): A class of multivariate failure time distributions. *Biometrika* 73, pp. 671-678.

Hougaard, P. (2000): Analysis of Multivariate Survival Data. Statistics for Biology and Health. Springer New York Berlin Heidelberg, 2000.

Hosmer, D.W., Lemeshow, S., May, S. (2008): Applied Survival Analysis. Regression Modelling of Time-to-Event Data. Second Edition John Wiley & Sons.

Hung W.S., Ho S.H. (2008): Survival Analysis for Unobserved Heterogeneity on Estimated Mortality in Taiwan. Economics Bulletin, Vol. No. 25 pp.1-10.

Hutson, J.M., Hasthorpe, S. (2005): Testicular descent and cryptorchismus: the state of the art in 2004. Journal of Pediatric Surgery 40, No. 2, pp.297-302

Huynh, T., Mollard, R., Trounson, A. (2002): Selected genetic factors associated with male infertility. Human Reproduction, Update, Vol.8, No.2 pp. 183-198.

Irvine, D.S. (1998): Epidemiology and aetiology of male infertility. Human Reproduction, 13, suppl 1, pp.33-44.

Jacobs, P.A.; Strong, J.A. (1959): A case of human intersexuality having a possible XXY sex-determining mechanism. Nature 183, pp.302-303.

Jensen, T.K., Jacobsen, R., Christensen, K., Nielsen, N.C., Bostofte, E. (2009): Good Semen Quality and Life Expectancy: A Cohort Study of 43,277 Men. *American Journal of Epidemiology*, Vol.170, No. 5 pp. 559-565.

Jonker, M.A., Boomsma, D.I., Bhulai, S., Ligthart, R.S.L., Posthuma, D., Van der Vaart, A.W. (2009): Gamma frailty model for linkage analysis with application to interval-censored migraine data. *Biostatistics*, 10, 1, pp.187-200.

Juul, S, Karmaus, W., Olsen, J. and the European Infertility and Subfecundity study group (1999): Regional differences in waiting time to pregnancy: pregnancy based surveys from Denmark, France, Germany, Italy and Sweden. *Human Reproduction* 14, pp.1250-1254.

Katzorke, T., Rickert-Föhring, M., Bielfeld, P., Kolodziej, F.B. (2008): Weibliche und männliche Fertilitätsstörungen und Risiken der assistierten Fertilisation In: Fritze, E., Mehrhoff, F. (Hrsg.) *Die ärztliche Begutachtung. Siebte, vollständige überarbeitete Auflage, Rechtsfragen, Funktionsprüfungen, Beurteilungen*, Steinkopff-Verlag Heidelberg.

Keck, C., Neulen, J., Behre, H.M., Breckwoldt, M. (2002): *Endokrinologie-Reproduktionsmedizin- Andrologie. 1. Band*, 2.Auflage, Thieme Verlag, Stuttgart, pp. 232-299.

Keyfitz, N (1978): Improving Life Expectancy: An Uphill Read Ahead. American Journal of Public Health, 10, pp.954-956.

Keyfitz, N., Caswell, H. (2005): Heterogeneity and Selection in Population Analysis. 3rd Edition Applied Mathematical Demography. Springer New York, pp. 477-493.

Klein, J.P., Moschberger, M.L. (2005): Survival Analysis Techniques for Censored and Truncated Data. Second Edition Springer Verlag, New York.

Kimura, M., Cherkas, L.F., Kato, B.S., Aviv, A. et al. (2008): Offspring's Leukozyte Telomere Length, Paternal Age, and Telomere Elongation in Sperm. PLoS Genetics, Vol.4, No.2.

Kirkwood, T.B.L. (1993): The disposable soma theory of aging: evidence and implications. Netherlands Journal of Zoology, 43, pp.359-363.

Kirkwood, T.B.L. (1999): Time of our lives. London: Weidenfeld and Nicolson.

Kitigawa, E.M., Hauser, P.M. (1973): Differential Mortality in the United States: A Study of Socioeconomic Epidemiology. Cambridge, MA: Havard University Press.

Kohler, U., Kreuter, F. (2006): Datenanalyse mit STATA, Allgemeine Konzepte der Datenanalyse und ihre praktische Anwendung. 2. vollständig überarbeitete und ergänzte Auflage, R. Oldenburg Verlag München Wien.

Kolon, T.F., Patel. R.P., Huff, D.S. (2004): Cryptorchidism: diagnosis, treatment, and long-term prognosis. Urologic Clinics North America, 31, No.3, pp. 469-80.

Konietzka, D., Kreyenfeld, M. (2005): Nichteheliche Mutterschaft und soziale Ungleichheit. Zur ökonomischen Differenzierung der Familienformen in Ost- und Westdeutschland. Kölner Zeitschrift für Soziologie und Sozialpsychologie (2005), Vol. 57 No. 1/März 2005 pp.32-61.

Könnecke, R., Wilke, S. (2002): Bewältigungsmuster bei ungewollt kinderlosen Männern, Ein Literaturüberblick und Ergebnisse einer qualitativen Studie. Reproduktionsmedizin (18) pp. 73-82.

Korpelainen, H. (2000): Fitness, reproduction and longevity among European aristocratic and rural Finnish families in the 1700s and 1800s. Proceeding of Royal Society London, 267, pp. 1765-1770.

Küpker, W., Schwinger E., Mennicke, K., Hiort, O., Bals-Pratsch, M., Ludwig, M., Schlegel P.N., Diedrich K. (2000): Genetik der männlichen Infertilität. Gynäkologe 2, pp.79-87.

Krause, W.(1993a): Die Bedeutung des „Routine-Spermiogramms“. Hautarzt 44, pp.269-274.

Krause, W., Weidner, W. (1999): Andrologie. Krankheit der männlichen Geschlechtsorgane. 3. völlig neu bearbeitete Auflage, Ferdinand Enke Verlag Stuttgart.

Lancaster, T. (1979): Econometric methods for the duration of unemployment. *Econometrica*, Vol.47, No.4 pp.939-956.

Le Bourg, E. Thon, B., Légaré, J., Desjardins B., Charbonneau H. (1993): Reproductive life of French-Canadians in the 17-18th centuries: a search for a trade-off between early fecundity and longevity. *Experimental Gerontology* 28, pp.217-32.

Le Bourg, E. (2001). A mini-review of the evolutionary theories of aging. Is it the time to accept them? *Demographic Research Volume 4, Article 1*, pp. 1-28, Max Planck Institute for Demographic Research.

Lee, T.Y., Sun, G.H., Chao, S.C. (2001): The effect of an infertility diagnosis on the distress, marital, and sexual satisfaction between husbands and wives in Taiwan. *Human Reproduction* 16 (8) pp. 1762-1767

Locatelli, I., Lichtenstein, P., Yashin A.I. (2003): The heritability of breast cancer: A Bayesian correlated Frailty Model applied to Swedish Twin Data. *Twin Research*, Volume 7, Number 2, pp.182-191.

Luzzi, G.A., O'Brien, T.S. (2001): Acute epididymitis. *British Journal of Urology International*, 87, No.8, pp. 747-755.

Lycett, J.E., Dunbar, R.I.M. , Volland, E. (2000): Longevity and the costs of reproduction in a historical human population. *The Royal Society*, 267, pp.31-35.

Lynch, S.M., Brown J.S., Harmsen, K.G. (2003): Black-White Differences in Mortality Compression and Deceleration and the Mortality Crossover Reconsidered. *Research on Aging* 25, pp.456-83.

Lin, D.Y. (1997): Non-Parametric inference for cumulative incidence functions in competing risk studies. *Statistics in Medicine*, Vol. 16, pp.901-910.

Liu, L., Wolfe, R.A., Huang, X. (2004): Shared frailty models for recurrent events and a terminal event. *Biometrics* 60 No. 3, pp. 747-56.

Manton, K.G., Stallard, E., Vaupel J.W. (1981): Methods for Comparing the Mortality Experience of Heterogeneous Populations, *Demography*, Vol. 18, No. 3, pp.389-410.

Manton, K.G., Poss, S.S., Wing, S. (1979): The Black/White Mortality Crossover: Investigation from the perspective of the components of aging. *The Gerontologist*, Vol 19, No.3., pp. 291-300.

Matsuura, M., Eguchi, S. (2005) Modelling Late Entry Bias in Survival Analysis. *Biometrics* 61, pp.559-566.

Medawar, P.B. (1952). *An unsolved problem in biology*. London: HK Lewis.

Mehrannia, T. (2007) The effect of cigarette smoking on semen quality of infertile men. *Pakistan Journal of Medical Sciences*, Vol. 23, No.5, pp. 717-719.

Meißner, A., Backhaus, B.O., Schierbaum, K.F., Müller, S.C., Haidl, G. (2007): Ätiologie und Therapie des Hodenhochstandes – Gesichtspunkte zur Optimierung der Fertilität. *Reproduktionsmedizin und Endokrinologie*, 4 (4), pp.177-184.

McGilchrist, C.A., Aisbett, C.W. (1991): Regression with frailty in survival analysis. *Biometrics* 47, pp.221-225.

Miksch, A., Herman, K., Trieschman, J., Roetz, A; Heiderhoff, M.; Laux, G., Rosemann, T; Szecsenyi, J. (2008): Geschlechtsspezifische Unterschiede in der Lebensqualität von Typ-2-Diabetikern mit und ohne DMP-Einschreibung. Gesundheitswesen, 70 (4), pp. 250-255.

Mirone V., Imbimbo, C., Bortolotti A., Di Cintio E., Colli, E., Landoni, M., Lavezzari, M., Parazzini, F. (2002): Cigarette Smoking as risk factor for erectile dysfunction: results from an Italian epidemiological study. European Urology , 41, pp.294-7.

Moreau, T., O'Quigley, J., Mesbah, M. (1985): A global goodness-of fit statistic for the proportional hazard model. Applied Statistics, 34, pp.212-218.

Moreau, T., O'Quigley, J., Lellouch, J. (1986): Schoenfeld's approach for testing the proportional hazard assumption. Biometrics, 73, pp. 513-515.

Mueller, U.O. (2004): Does Late Reproduction extend the Life Span? Findings from European Royalty. Population and Development Review 30, pp. 449-466.

Müller, H.G., Chiou, J.M., Carey, J.R., Wang, J.L. (2002): Fertility and Lifespan: Late Children enhance female Longevity. Journal of Gerontology Series A. Biological and Medical Sciences, Vol. 57, No.5, pp. 202-206.

Nagle, H.M.: Editorial Comment (2004): Varicocele: where, why and, if so, how? The Journal of Urology, 172., pp. 1239-1240.

Nallella, K., Sharma, R., Aziz, N., Agarwal, A. (2006): Significance of sperm characteristics in the evaluation of male infertility. Fertility and Sterility, Vol.85, No. 3, pp. 629-634.

Nair, V.N. (1984): Confidence bands for survival with censored data. Technometrics, 26, pp.265-275.

Ng, K.K., Donat, R., Chan, L., Lalak, A., Di Pierro, I., Handelsman, D.J. (2004): Sperm output of older men. Human Reproduction, Vol. 19, No.8, pp.1811-1815.

Ng'andu, N.H.(1997): An empirical comparison of statistical test for assessing the proportional hazards assumption of Cox's Model. Statistics in Medicine, Vol.16, pp.611-626.

Nieschlag, E. (1993): Care for infertile male (current therapie). Clinical Endocrinology 38, pp.123-133.

Nieschlag E., Behre H.M. (2009): Andrologie. Grundlagen und Klinik der reproduktiven Gesundheit des Mannes. 3. Auflage Berlin Springer, 2003.

Oakes, A. (1982): A Model for Association in Bivariate Survival Data. Journal of the Royal Statistical Society. Series B (Methodological), 44 (3), pp. 414-422.

Omori, Y., Johnson, R.A.(1993): The influence of random effects on the unconditional hazard rate and survival functions. Biometrika, 80, pp.910-4.

O'Quigley, J., Pessione, F. (1989): Score tests for homogeneity of regression effect in the proportional hazard model. Biometrics, 45; pp. 135-144.

Ozgur, K, Isikoglu, M., Seleker, M., Donmez, L. (2005): Semen quality of smoking and non-smoking men in fertile couples in a Turkish population. Archives of Gynecology and Obstetrics, 271, pp.109-112.

Pan Wei, Chappell, R. (2002): Estimation in the Cox Proportional Hazards Models with Left-Truncated and Interval-Censored Data. Biometrics, Vol... 58, No.1, pp.64-70.

Pasqualotto, F.F., Umezu, F.M., Salvador M., Borges E., Sobreirio B.P, Pasqualotto, E.B.(2008): Effect of cigarette smoking on antioxidant level and presence of leukocytospermia in infertile men: a prospective study. Fertility and Sterility, 90, pp. 278-83.

Pérez-Marín, A.M. (2008): Empirical comparison between the Nelson-Aalen Estimator and the Naïve Local Constant Estimator. *Statistics and Operation Research Transaction*, 1696-2281, Vol.32, No.1, pp.67-76.

Petersen, L., Sørensen, T.I., Andersen, P.K. (2010): A shared frailty model for case-cohort samples: parent and offspring relations in an adoption study. *Statistics in Medicine*, 29 (7-8) pp.924-31.

Pike, M.C. (1966): A method of analysis of a certain class of experiments in carcinogenesis. *Biometrics* 22, pp.142-161.

Preston, S.H., Elo, I.T., Rosenwaikke I., Hill M.E. (1996): African-American Mortality at Older Ages: Results of a Matching Study. *Demography*, Vol. 33, No.2, pp.193-209.

Ripatti, S., Palmgren, J. (2000): Estimation of Multivariate Frailty Models Using Partial Likelihood. *Biometrics*, 56(4), pp. 1016-1022.

Rabe-Hesketh, S., Skrondal A. (2008): *Multilevel and Longitudinal Modeling Using Stata*.

2nd Edition, Stata Press Publication StataCorp LP , College Station Texas, 2008.

Robert, C. (2003): Etude et modelisation du fonctionnement d'un convert de blé attaqué par le complexe parasitaire *Puccinia triticina*-*Mycosphaerella graminicola*. Ph.D. thesis, Institut National Agronomique Paris-Grignon.

Rondeau, V., Commenges, D., Joly, P. (2003): Maximum Penalized Likelihood Estimation in a Gamma-Frailty-Model. *Lifetime Data Analysis* 9, pp.139-153.

Roger, R.G. (1992): Living and Dying in the U.S.A.: Sociodemographic Determinants of Death Among Blacks and Whites. *Demography* 33, pp. 193-209.

Rosenberger, W.F., Sessaiger, P. (1997): Adaptive survival trails. *Journal of Biopharmaceutical Statistics* No. 7, Vol.4, pp.617-624.

Rosengren, A., Orth-Gomer, K., Wedel, H. und Wilhelsen, L. (1993): Stressful live events, social support, and mortality in men born in 1933. *British Medical Journal*, 307, pp. 1102-1105.

Rothman, K.J., Greenland S.(1998): *Modern Epidemiology*. Second Edition. Philadelphia, Pa.: Lippincott-Raven, 1998.

Runnebaum, B., Rabe, T. (1994): *Fortpflanzungsmedizin. Gynäkologische Endokrinologie und Fortpflanzungsmedizin*. Springer-Verlag Band 2.

Schill, W.B., Comhaire, F.H., Hargreave, T.B. (2006): Andrology for the clinician, Springer Verlag Berlin Heidelberg

Siler, W. (1979): A Competing-Risk Model for Animal Mortality. Ecology, Vol. 60, No.4, pp.750-757.

Schmidt, R.F., Lang, F., Thews G. (2007): Physiologie des Menschen, mit Pathophysiologie, 30. neu bearb. und aktualisierte Auflage, Springer-Verlag Heidelberg, Berlin.

Schmitt, C. (2004): Kinderlose Männer in Deutschland – Eine sozialstrukturelle Bestimmung auf Basis des Sozio-ökonomischen Panels (SOEP). Kurzexpertise DIW Berlin.

Schmitt, C., Winkelmann, U. (2005): Wer bleibt kinderlos? Sozialstrukturelle Daten zur Kinderlosigkeit von Frauen und Männern. DIW Berlin Discussion Papers 473.

Schoenfeld, D. (1982): Partial residuals for the proportional hazard regression model, Biometrika, 69, pp. 239-241.

Schoenfeld, D. (1980): Chi-squared goodness-of-fit tests for the proportional hazards regression models, *Biometrika*, 67, pp. 145-153.

Schumacher (1990): Evaluation of nonproportional treatment effects in cancer clinical trials. *Cancer Investigation*, 8, pp.91-98.

Sear, R. (2007): The impact of reproduction on Gambian Women: Does Controlling for Phenotypic Quality reveal costs of reproduction. *American Journal of physical anthropology* 132, pp.632-641

Semenchenko, G.V., Yashin, A.I., Johnson, T.E., Cypser J.W. New Modells for Survival Analysis of Experimental Data. In: Agente, J-L.: *Advances in Statistical Methods for the Health Sciences, Application to Cancer and AIDS studies, Genome Sequence Analysis and Survival Analysis*, pp. 179-191, (2007) Birkhäuser Boston Basel Berlin.

Seydel, H. (2010): Überlebenszeit und Todesursachen fertiler und subfertiler Männer. Dissertation Philipps-Universität Marburg

Sharpe, RM., Skakkebaek, N.E. (1993): Are oestrogens involved in falling sperm counts and disorders of the male reproductive tract? *Lancet* 341 (8857), pp.1392-1395.

Shye, D., Mullooly, J.P., Freeborn, D.K., Pope, C.R. (1995): Gender differences in the relationship between social network support and mortality: A longitudinal study of an elderly cohort. *Social Science and Medicine*, Vol. 41, Issue 7, pp. 935-947

Silbernagel S, Despopoulus A. Taschenatlas der Physiologie, 5. Auflage, Thieme Verlage, Stuttgart, New York, München, 2001; pp. 306-307.

Siler, W. (1979): A competing risk model for animal mortality. *Ecology* 60 (4), pp. 750-757.

Smith, K.R., Mineau, G.P., Bean, L.L. (2002): Fertility and post-reproductive longevity. *Social Biology* 49, pp.185-205.

Sonnenschein, C., Soto A.M. (1998): An updated review of environmental estrogen and androgen mimics and antagonists. *Journal of Steroid Biochemistry and Molecular Biology*. 65 (1-6), pp.143-150.

Soubeyrand, S., Sache, I., Lannou, C., Chadœuf (2007): A Frailty Model to Assess Plant Disease Spread from Individual Count Data. *Journal of Data Science* 5, pp.67-83.

Springer, A., Horcher, E. (2009): Hypospadie, Risikofaktoren - chirurgische Therapie – psychosoziale Aspekte. Pädiatrie und Pädologie, Volume 44, No.1 , pp.24-27.

Sun, J. (2006): The Statistical Analysis of interval-censored Failure Time Data, Statistics for Biology and Health. 2006 Springer Science + Business Media, Inc. New York.

Stata Corp. (2008): Stata Statistical Software: Release 10.0. College Station, TX, StataCorp LP.

Takase, B., Etsuda H., Matsushima Y., Ayaori, M., Kusano, H., Hamabe, A., Uehata A., Ohsuzu F., Ishihara M., Kurita, A. (2004): Effect of chronic oral supplementation with vitamins on the endothelial function in chronic smokers. Angiology, 55 (6): 653.

Therneau, T.M., Grambsch, P.M., and Fleming T.R.(1990): Martingale-based residuals for survival models. Biometrika, 77, pp.147-160.

Therneau, T.M., Grambsch P.M. (2000): Modelling Survival Data: Extending the Cox Model.
New York:Springer.

Thernau, T.M., Grambsch, P.M. (2003): Penalized Survival Models and Frailty. *Journal of Computational and Graphical Statistics*, 12(1), pp.156-175.

Thomas, F., Teriokhin, T., Renaud, F., De Meeûs, T. & Guégan, J.F. (2000): Human longevity at the costs of reproductive success: evidence from global data. *Journal of Evolutionary Biology* 13, pp. 409-414.

Trimbos, B., Timmers, P., Pecorelli, S., Coens, C., Ven, K., van der Burg M., Casado, A. (2010): Surgical Staging and Treatment of Early Ovarian Cancer: Long-term Analysis from a Randomized Trail. *Journal of the National Cancer Institute*, Vol.102, No.13, pp.982-986.

Trivon, I.M. (2010): Analytic Expression for Life Expectancy in Gamma-Gompertz Mortality –Setting. MPIDR Working Paper, WP2010-020.

Van Weert, J.M., Repping, S., Van Voorhis, B.J., van der Veer, F., Bossuyt, P.M.M, Mol, B.W.J.(2004): Performance of the postwash total motile sperm count as a predictor of pregnancy at the time of intrauterine insemination: A meta-analysis. *Fertility and Sterility*, Vol.82, No.3, pp.612-620.

Van Weert, J.M., Repping, S., van den Steeg, J.W., Streuers, P., van der Veen, F., Mol, B.W. (2008): A prediction model for ongoing pregnancy after in vitro fertilization in couples with male subfertility. *Journal of Reproductive Medicine* 53, 4, pp. 250-256.

Vaupel, J.W., Manton, K., and Stallard, E. (1979): The impact of heterogeneity in individual frailty on dynamics of mortality. *Demography*, 16, pp. 439-454.

Virtanen, H.E., Toppari, J. (2008) Epidemiology and pathogenesis of cryptorchidism. *Human Reproduction Update*, 14, pp.49-58.

Vittinghoff, E., Glidden, D.V., Shiboski, S.C., McCulloch, C.E. (2005): *Regression Methods in Biostatistics, Linear, Logistic, Survival and Repeated Measures Models*. Springer Science+Business Media, New York.

Vogt, P.H., Edelmann A., Kirsch, S. et al. (1996): Human Y chromosome azoospermia factor (AZF) mapped to different subregions in Yq11. *Human Molecular Genetics*, 5, pp. 933-943.

Westendorp, R.G.J. & Kirkwood T.B.I. (1998): Human longevity at the cost of reproductive success. *Nature* 396, pp. 743-746.

Wienke, A., Christensen K., Skytthe, A., Yashin, A.I. (2002): Genetic analysis of cause of death in a mixture model with bivariate lifetime data. *Statistical Modelling*, 2(2), pp. 89-102

Wienke, A., Arbee, K., Locatelli, I., Yashin, A.I. (2003): A simulation study of different correlated frailty models and estimation strategies. Technical report, MPIDR Working Paper WP 2003.

Wienke, A. (2004): Die Vererbbarkeit von Todesursachen: ein „correlated frailty“-Modell angewandt auf dänische Zwillinge, geboren 1870-1930. In: Scholz, R., Flöthmann, J. *Lebenserwartung und Mortalität. Materialien zu Bevölkerungswissenschaften* Heft 111, Wiesbaden, 81-89.

Wienke, A., Lichtenstein, P., Czene, K. and Yashin, A.I.. The role of correlated Frailty Models in studies of Human Health, Ageing and Longevity. In: Agote, J-L.: *Advances in Statistical Methods for the Health Sciences, Application to Cancer and AIDS studies, Genome Sequence Analysis and Survival Analysis*, pp. 151-161, (2007) Birkhäuser Boston Basel Berlin.

Wienke, A., Ripatti, S., Palmgren, J., Yashin, A. (2010): A bivariate survival model with compound poisson frailty, *Statistics in Medicine* 29, pp. 275-283

Winnet, A., Sasieni, P. (2001): A note on scaled Schoenfeld residuals for the proportional hazard model, *Biometrika*, 88, 2, pp.565-571

Wintrebert, C.M.A., Zwinderman, A.H., Cam, E., Pradel, R., van Houwelingen, J.C. (2005):

Joint modelling of breeding and survival in the kittiwake using frailty models. *Ecology Modelling*, 181, pp.203-213.

Willcox, D.G., Willcox, B.J., Hsueh, W.C., Suzuki, N. (2006): Genetic determinants of exceptional human longevity: insights from the Okinawa Centenarian Study. *AGE* Volume 28, No. 4/ December 2006.

Williams, G.C. (1957): Pleiotrophy, natural selection and the evolution of senescence. *Evolution* 11: pp. 398-411.

Wise, G.J. und Marella, V.K. (2003): Genitourinary manifestations of tuberculosis. *Urologic Clinics of North America*, 30, No.1, pp.111-121.

World Health Organisation (1999): WHO laboratory manual for the examination of human semen and sperm-cervical mucus interaction, 4 edn. Cambridge University Press, Cambridge.

World Health Organisation (2010): Laboratory manual for the examination and processing of human semen. 5. edn. Geneva, Switzerland.

Xu, R. (2004): Proportional Hazards Mixed Models: A review with applications to twin models. *Methodološki zvezki*, Vol. 1, No1, pp.205-212.

Yashin, A.I., Vaupel, J.W., Iachine, I.A. (1995a): Correlated individual frailty: An advantageous approach to survival analysis of bivariate data. *Mathematical Population Studies*, Vol.5(2), pp.145-159.

Yashin, A.I., Iachine, I.A. (1995b): Genetic analysis of durations: correlated frailty model applied to survival of Danish twins. *Genetic Epidemiology* 12(5), pp.529-538.

Yashin, A.I., Arbeey, K.G., Akushevich, I., Kulminski, A., Akushevich, L., Ukraintseva, S.V. (2008): Model of hidden heterogeneity in longitudinal data. *Journal of Theoretical Biology*, 73(1) 1-10.

Yasuhiro, O., Johnson R.A. (1993): The influence of random effects on the unconditional hazard rate and survival functions. *Biometrika*, 80, 4, pp.910-4.

Zavos, P.M., Zarmakoupis-Zavos, P.N. (1999): Impact of cigarette smoking on human reproduction: Its effects on male and female fecundity. Technology, Vol. 6, pp9-16.

Ying, G.S., Heitjan, D.F.(2008): Weibull prediction of event time in randomized clinical trials. Pharmaceutical Statistics No.7, pp.107-120

Zavos P.M., Correa J.R., Antypas, S., Zarmakoupis-Zavos P.N., Zarmakoupis C.N. (1998): Effects of seminal plasma from cigarette smokers and the risk of male factor subfertility:minor association between cotinine in seminal plasma and semem morphology. Fertility and Sterility , 69 ,pp.425-9.

Zenzes, M.T., Bielecki, R., Reed, T.E. (1999a): Detection of benzo(a)pyrene diol epoxide-DNA damage aducts in sperm of men exposed to cigarette smoke. Fertility and Sterility, 72, pp.330-335.

Zenzes, M.T., Bielecki, R., Reed, T.E. (1999b): Detection of benzo(a)pyrene diol epoxide-DNA aducts in embryos from smoking couples: evidence for transmission by spermatozoa. Molecular Human Reproduction, 5 pp.125-131.

Zitzmann, M., Nieschlag, E. (2004): Rauchende Männer- ein Fertilitätsrisiko. Journal für Reproduktionsmedizin und Endokrinologie ; 1(1), pp9-12.

Zhong, X., Li, H. (2002): An additive genetic gamma frailty model for two-locus linkage analysis using sibship age of onset data. Statistical Application in Genetic and Molecular Biology, Vol.1, No.1, Berkley Electronic Press 2002.

Onlinequellen:

Albers, P., Albrecht, W., Algaba, F., Bokemeyer, C. und Andere (2006): Guidelines on Testicular Cancer. European Association of Urology.

www.uroltoday.com/prod/pdf/eau/09_Testicular_Cancer.pdf

Stand: 01.11.2009

Albrecht, M.A. (2008): Epidemiology, clinical manifestations, diagnosis and management of mumps.

www.utdol.com/patients/content/topic.do?topicKey=~KcFMMOAlxVebIN7

Stand: 23.05.2008

Box-Steffensmeier, J.M., Zorn C. (1999): Modeling Heterogeneity in Duration Models. 1999 Summer Meeting of Political Methodology Society, Texas A&M University, College Station, TX.

<http://polmeth.wustl.edu/retrieve.php?id=235>

Stand: 28.07.2008

Center For The Evaluation of Risks To Human Reproduction (CERHR)
(2006):

NTP-CERHR Monograph on the Potential Human Reproductive and
Development Effect of Di(2-Ethylhexyl) Phthalate (DEHP), National
Toxicology Program US Department of Health and Human Services, NIH
Publication No. 6-4476.

<http://cerhr.niehs.nih.gov/chemicals/dehp/DEHP-Monograph.pdf>

Stand: 06.10.2009

Chen, X. (2008): Score Test of Proportionality Assumption for Cox Models.
Statistical Consulting Group UCLA, Los Angeles, CA.

www.wuss.org/proceedings08/08wuss%20Proceedings/papers/an1/anl05.pdf

Stand: 05.08.2009

Christensen, K., Johnson Thomas E., Vaupel J.W. (2006): The quest for
genetic determinants of human longevity: challenges and insights.

<http://www.nature.com/reviews/genetics> Stand: 14.08.2009

DocCheck (2008): Urogenitaltuberkulose.

<http://www.flexikon.doccheck.com/Urogenitaltuberkulose.html>

Stand 06.08.2008

Dohle, G.R., Weidner, W., Jungwirth, A., Colpi, G., Papp, G., Pomerol; J., Hargreave, T.B. (2004): Guidelines on male infertility, European Association of Urology.

www.urotoday.com/prod/pdf/eau/22891_Male_Infertility.pdf

Stand: 06.08.2008

Dohle, G.R., Jungwirth, A., Colpi, G., Giwercman, A., Diemer, T., Hargreave, T.B. (2007)

Guidelines on male infertility. European Association of Urology, Update.

www.roweb.org/fileadmin/user_upload/guidelines/13_Male_Infertility_2007.pdf

Stand: 06.08.2008

Dorn, K. (2007): Einflussfaktoren von Seiten des Mannes auf den Erfolg einer in-vitro-Fertilisation mit intracytoplasmatischer Spermieninjektion (ICSI). Klinik für Urologie und Kinderurologie der Universität Ulm.

http://deposit.ddb.de/cgi-dinldokserv?=997456248&dok_var=d1&dok_ext=pdf&filename=997456248.pdf

Stand: 17.09.2008

Drefahl, S. (2008): How does the age gap between partners affect their survival? Paper presented at Annual Meeting of the Population Association of America 2008, New Orleans, LA.

<http://paa2008.princeton.edu/download.aspx?submissionId=80997>.

Stand: 17.08.2008

Engelman, M., Canudas-Romo, V., Agree, E. (2009): Variation and Vulnerability: Unobserved Heterogeneity in the Course of Mortality Transitions

<http://epc2010.princeton.edu/download.aspx?submissionId=100399>

Stand 24.05.2010

Elwert, F., Wrigley-Field, E. (2009): Earlier Race Crossover in Mortality in the United States Center for Demography and Ecology, University of Wisconsin-Madison. Presented as Poster at Annual Meeting of the Population Association of America PAA 2009, Detroit, MI.

<http://paa2009.princeton.edu/download.aspx?submissionId=91905> Stand:

04.06.2009

Eskenazi, E., Wyrobek, A.J., Slotter, E. , Kidd, S.A., Moore, L. Young, SS., Moore D. (2001): Decrease in Human Semen quality with Age among healthy men. Second International Meeting on Male Medicated Development Toxicity, Montreal Canada, June 20-23, 2001.

<http://www.osti.gov/bridge/purl.cover.jsp;jsessionid=4A7CBB94F41E992DD5C9FDBCF727E3?purl=/802878?purl=/802878-UQrCl8/native/>

Stand: 10.02.2009

Fox, J. (2002) Cox Proportional-Hazards Regression for Survival Data. Appendix to An R and S-Plus Companion to Applied Regression.

<http://cran.r-project.org/doc/contrib/Fox-companion/appendix-cox-regression.pdf>

Stand: 10.03.2009

Gagnon, A., Mazan, R., Desjardins B., Smith K.R. (2005): Post-reproductive longevity in a natural fertility population. Discussion Paper no.05-04. Population Studies Centre University of Western Ontario, London.

<http://www.iussp.org/members/restricted/publication/Utah05/Gagnon.pdf>

Stand: 02.06.2009

Gagnon, A., Smith, K.R., Tremblay, M., Vezina, H., Pare, P.P., Desjardins, B. (2009):

Is there a trade-off between fertility and longevity? A comparative study of women from three large historical databases accounting for mortality selection. Paper presented at Annual Meeting of the Population Association of America PAA 2009, Detroit, MI.

<http://www.paa2009.princeton.edu/download.aspx?submissionId=90349>

Stand: 27.05.2009

Ghosh, S.K., Ghosal S. (2006): Semiparametric Accelerated Failure Time Models for Censored Data, Department of Statistics, North Carolina State University Raleigh.

<http://www4.stat.nseu.edu/~sgghosal/papers/PMR.pdf>

Stand: 14.09.2008

Groos, S. (2006): Lebenszeit-Mortalität von Männern mit normalen und subnormalen Spermienkonzentrationen. Philipps-Universität Marburg.

http://deposit.ddb.de/cgi-bin/dokserv?idn=979333261&dok_var=d1&dok_ext=pdf&filename=979333261.pdf

Stand 15.07.2007

Harrell, F.E. (1986): " The PHGLM procedure", SAS supplemental Library User's Guide, Version 5 Edition, SAS Institute, Cary, N.C.

[http://openlibrary.org/b/OL3831315M/SAS supplemental Library User Guide](http://openlibrary.org/b/OL3831315M/SAS_supplemental_Library_User_Guide)

Stand: 11.03.2009

Holm, A., Jæger M., Pedersen M. (2008): Unobserved heterogeneity in the binary logit model with cross-sectional data and short panels: a finite mixture approach. The Danish national Centre for social research, 16:2008 working paper.

<http://www.sfi.de>

Stand 14.02.2009

Hampel, F. (1998): On the Foundations of Statistics: A frequentist approach. Seminar für Statistik Eidgenössische Technische Hochschule (ETH) Zürich, Research Report No.85.

<ftp://ftp.stat.math.ethz.ch/Research-Reports/85.pdf>

Stand: 02.11.2009

Jakobi, H. (1999): Rückgang der Spermienqualität in Deutschland und Europa. Ein Bericht des Institutes für angewandte Toxikologie und Umwelthygiene an der Universität Oldenburg.

http://www.freiheitist selbstbestimmtesleben.de/Rueckgang_Spermienqualitaet_Studie_WMF.pdf

Stand 5.03.2008

Keele L. (2007): Cross-Validation Tests for Frailty Models. Ohio State University.

<http://www.polisci.ohio-state.edu/faculty/lkeele/keele.html>

Stand: 28.01.2009

Krause, W. (1993b): Einfluß von Ejakulat-Parametern auf die Fertilitätsprognose, Jahrestagung der Deutschen Gesellschaft für Andrologie, Hamburg 24/25.09.1993.

<http://www.andrologie.de>

Stand: 09.10.2009

Mair P., Hudec, M. (2008): Session Clustering Using Mixture of Proportional Statistics and Mathematics, Universität Wien, Research Report Series, Report 63.

<http://epub.wu-wien.ac.at>

Stand: 06.08.2008

Mair P., Hudec, M. (2007): mixPHM: R package for mixture of proportional hazard models. <http://cran.R-project.org>

Stand: 06.08.2008

Medical Tribune (2009): Ovarialkazinome wachsen symptomlos. Meist erst spät diagnostiziert.

<http://www.medical-tribune-at/dynasite.cfm?dsmid=77069&dspaid=644484>

Stand: 02.04.2009

MedHost (2008): Nebenhodenentzündung.

<http://www.medhost.de/Sexualität-partnerschaft/sexuell-uebertragbare-krankheiten/nebenhodenentzuendung.html>.

Stand: 07.08.2008

Mitteldorf, Joshua (2009): Female Fertility and Longevity. Department of Ecology and Evolutionary Biology. University of Arizona Tucson, AZ.

<http://arxiv.org/ftp/arxiv/papers/0904/0904.1815.pdf>

Stand: 21.01.2009

National Center for Health Statistics (2003): Health, United States. Hayattsville, MD: U.S. Department of Health and Human Services.

<http://www.cdc.gov/nchs/> Stand: 07.05.2008

Petersen, E. (2005): Phthalate- Minimierungsstrategien im Gesundheitswesen.

<http://www.gewhessen.de/fileadmin/uploads/dokumente/pthalate.pdf>

Stand: 23.07.2009

Persson I., Astra-Zenecar, R., Söndertälje, D., Khamis, H. (2007): A Comparison of graphical methods for assessing the proportional hazard assumption in the Cox Model, Festschrift in honor of Distinguished Professor Mir Masoom Ali, pp.23-43.

http://dumbach.iweb.bus.edu/AliConference/papers/File04_Khamis.doc

Stand: 23.07.2009

Rodriguez, G. (2001): Unobserved Heterogeneity. The Office of Population Research at Princeton University.

http://data.princeton.edu/pop505a/Unobserved_Heterogeneity.pdf

Stand: 07.08.2008

Schräder, G. (2004): Der Einfluß des Rauchens auf das Ergebnis der Assistierten Reproduktion.

<http://miami.uni-muenster.de/servlets/DocumentServlet?id=1980.pdf>

Stand 17.01.2008

Shaban A., Mostafa, A.A. (2003): Shared Frailty Survival Analysis. Using Semiparametric Bayesian Method. University of Cairo.

<http://www.interstat.statjournal.net/YEAR/2005/articles/0511002.pdf>

Stand: 27.01.2009

Silber, S.J. (2003): The Genetics of male infertility. Evolution of the the X and Y chromosome and transmission of male fertility to future generations. Infertility Center of St. Louis, MO.

www.infertile.com/pdf_files/intheview/sci/xyevolution.pdf

Stand: 03.09.2008

Soares, S (2009): Cigarette Smoking and Fertility. Reproductive Biology Insights 2009;2 pp.39-46.

<http://www.la-press.com>

Stand: 08.10.2009

Sommer, F. (2004): Die Einflüsse des Fahrradfahrens auf die Sexualität- Teil2: Hat Radsport einen Einfluß auf die Spermaqualität? Blickpunkt DER MANN, 2 (1), pp. 33-36.

<http://www.kup.at/dermann>

Stand: 12.11.2009

Statistisches Bundesamt (2006): Generationen-Sterbetafeln für Deutschland. Modellrechnung A, Generationensterblichkeit mit vollständiger Kriegsterblichkeit 1871-2004, Trend V1(Entwicklung seit 1871). Männliche Jahrgänge für das frühere Bundesgebiet.

http://www.destatis.de/jetspeed/portal/cms/Sites/destatis/Internet/DE/Navigat ion/Publikationen/Fachveroeffentlichungen/Bevoelkerung.templateId=render Print.psml_nnn=true

Stand: 08.02.2008

Statistisches Bundesamt (2008): Zahl der Eheschließungen ging im Jahr 2007 um 2% zurück.

Pressemitteilung Nr. 317 vom 28.08.2008.

http://www.destatis.de/jetseed/portal/cms/Sites/destatis/Internet/DE/Presse/p m/2008/08/PD08_317_12631.templateId=renderPrint.psml

Statistisches Bundesamt (2007): Geburten in Deutschland.

<http://www.destatis.de/jetseed/portal/cms/Sites/destatis/Internet/DE/Content/ Publikationen/Fachveroeffentlichung/Bevoelkerung/BroschuereGeb>

Stand: 28.08.2008

Techniker Krankenkasse online (2008): Varikozele.

<http://www.tk-online.de/tk/krankheiten-a-z/krankheiten-v/varikozele/31360>

Stand: 06.08.2008

Tomazella, V.C.D., Martins, C.B., Bernado, J.M. (2006): Inference on the univariate frailty model. A Bayesian reference analysis approach. Universidade Federal de Sao Carlos, SP-Brazil.
<http://adsabs.harvard.edu/abs/2008AIPC.1073.340T>

Stand: 28.06.2009

Tuljapurkar, S., Edwards R.D. (2011): Variance in death and its implication for modelling and forecasting mortality. Demographic Research Volume 24, Article 21, pp.497-526.
<http://www.demographic-research.org/Volumes/Vol24/21>

Stand: 14.06.2011

Turek, P. (2008): The genetics of male infertility. The Turek Clinic, Men's Reproductive Health Specialists.
<http://www.theturekclinic.com/PDF/turek-clinic-male-infertility-genetics.pdf>

Stand: 13.03.2009

u.V. (2008): Krankheitsbilder, Unfruchtbarkeit (Infertilität) und Andrologie, unerfüllter Kinderwunsch (Unfruchtbarkeit).
[www.urologie-remscheid.de/krankheitsbilder,unfruchtbarkeit-infertilität- und andrologie, 3.htm](http://www.urologie-remscheid.de/krankheitsbilder,unfruchtbarkeit-infertilität-und-andrologie,3.htm)

Stand 15.08.2008

Vaupel, J.W., Rau, R., Camarda, C.G., von Kistowski, K.G. (2006): Can Heterogeneity of population explain differences in mortality? Center for Retirement Research at Boston College. CPR WP 2006-10.
<http://www.bc.edu/crr>

Stand: 02.06.2009

Vaupel, J.W., Zhang, X., Engelman, M. (2009): The Impact of Disparities on the Age Trajectory of Mortality. Paper presented at Annual Meeting of the Population Association of America 2009, Detroit, MI.
<http://paa2009.princeton.edu/download.aspx?submissionId=91046>

Stand: 02.06.2009

Vermunt J.K. (2002): A general latent class approach to unobserved heterogeneity in the analysis of event history data.

<http://splitwww.uvt.nl/~vermunt/hagenaars2002apdf>

Stand: 11.08.2008

Westerman, R., Pyka K., Seydel H., Krause, W., Mueller U.O. (2009): Fertility Biomarkers predict differential survival in a long-term follow-up study of male fertility: Capturing and treating heterogeneity. Paper presented at Annual Meeting of the Population Association of America 2009, Detroit, MI.

<http://paa2009.princeton.edu/download.aspx?submissionId=90082>

Stand: 02.06.2009

Wienke, A. (2006): Frailty Models in Survival Analysis. Habilitation an der Medizinischen Fakultät der Martin-Luther-Universität Halle-Wittenberg.

<http://sundoc.bibliothek.uni-halle.de/habil-online/07/07H056/index.htm>

Stand 09.11.2008

Woutersen, T. (2007): Testing for Heterogeneity in Duration Models. A semiparametric Test. John Hopkins University Baltimore, MD.

<http://www.cemap.ac.uk/lancaster/woutersen.pdf>

Stand 19.05.2009

Yashin, A.I., Iachine, I.A., Begun, A.Z. und Vaupel, J.W. (2001): Hidden Frailty: Myths and Reality. Discussion Paper, Max Planck Institute for Demographic Research.

http://www.demogr.mpg.de/publications/file/1993_1121953571_1_RR34.pdf

Stand: 02.12.2009

Abbildung A1: Bimodale Population-Hazards und Hazard-Funktion separat für den jeweiligen Peak, Haugen, et. al. 2009

